Size-dependent tax enforcement and compliance: global evidence*

Pierre Bachas†, Anders Jensen ‡

January 2017

Abstract

In this paper, we provide identified estimates of how firm size impacts tax enforcement and compliance, using data from 125,000 firms in 140 countries. The magnitude of these size-gradients determine the quantitative importance of size-dependent policies in resource misallocation. The identification strategy uses the relative ranking of industries' average firm-size in the US as an instrument for the relative ranking of the same industries' size in a developing country. The first stage coefficient is robustly estimated and constant over levels of development. We find large and positive effects of firm size on tax enforcement and tax compliance and a negative effect on informal payments (bribes). These relations are non-linear in the size-distribution: tax enforcement and compliance are U-shaped, while formal tax payments appear to substitute informal payments as firms grow. Finally, the size gradient in enforcement is strongest at lower levels of development and zero in high income countries.

*We are grateful to Alan Auerbach, Tim Besley, Michael Best, Natalie Cox, Tom Cunningham, Simon Galle, Roger Gordon, Michael Keen, Henrik Kleven, Camille Landais, Florian Misch, Ted Miguel, Torsten Persson, Andres Rodriguez-Clare, Emmanuel Saez, Johannes Spinnewijn, Munir Squires, Owen Zidar and seminar participants at the LSE, the 69th IIPF Congress and the 2014 ZEW Public Finance conference for valuable comments. This paper builds on earlier findings from 2014 circulated under the title "An Empirical Test of Information Trails: Do Firm Finance and Size Matter for Tax Collection?" Bachas gratefully acknowledges financial support from the Center for Equitable Growth. Jensen gratefully acknowledges financial support from the Peter G. Foundation Grant No.16017 and the ESRC.

†Princeton University and World Bank Research

‡Harvard Kennedy School and NBER

1 Introduction

An influential literature explains cross-country differences in income and productivity through the misallocation of resources across firms (Restuccia and Rogerson 2008, Hsieh and Klenow 2009). One way resource misallocation occurs is through government policies that depend on firm size; for example size-dependent labor regulations, tax rates or accounting requirements could discourage firm growth and affect the firm size distribution (Gollin 1995, Guner et al. 2008). The magnitude of the distortions from size-dependent policies is the product of firms' size elasticity with respect to the policy variable and the slope of the policy variable with respect to firm size. This paper focuses on estimating the latter term, which we name the size-gradient. Estimating policy size-gradients in a systematic manner is challenging. First, while some policies have explicit size thresholds, other size-dependent policies follow a discretionary implementation and are therefore hard to observe. Second, data availability limit researchers' ability to characterize size-gradients across countries and in the full firm-size distribution.

In this paper, we provide identified evidence of size-gradients across a large set of countries and across the firm-size distribution. To this end we focus on a specific policy, *tax enforcement*. Tax enforcement may depend on firm size for a number of reasons, such as revenue potential or audit cost-effectiveness being concentrated amongst specific size-groups (Kleven et al. 2016). Tax enforcement is relevant for the study of policy-driven distortions for at least three reasons. First, tax enforcement directly impacts firms' effective tax rates, generating a wedge across firms of different size. Second, many firms report facing significant costs when dealing with the tax administration (World Bank 2017).[1] Third, reliance on size-dependent policies for tax enforcement has increased over time, as international institutions encouraged tax administrations to segment taxpayers (Kanbur and Keen 2014). To illustrate this trend, Figure 1 shows that over the past 20 years, more than 70 countries adopted size-dependent enforcement units for large taxpayers. While large taxpayer units could imply that enforcement is stringent at the top of the firm-size distribution, countries have also adopted enforcement policies targeted at small and medium firms.[2] Therefore how tax inspection and compliance vary over the full firm-size distribution is an empirical question.

[1] The World Bank 2017 Doing Business publication reports that "*on average taxpayers spend 25 hours complying with the requirements of an auditor, and go through several rounds of interactions during 10.6 weeks.*"

[2] Many countries implemented small and medium taxpayer offices alongside the large taxpayer office. In these small and medium offices, audit algorithms are tailored to the avoidance and evasion risks specific to smaller firms.

In our analysis we show how tax enforcement and tax compliance vary across the full firm-size distribution, in countries around the world. We use the comprehensive World Bank Enterprise Surveys (WBES), which contain firm-level data on self-reported tax inspection, sales tax compliance and informal tax payments,[3] for a sample of 125,000 firms in 140 countries. Exploiting arguably exogenous variation in an industry's optimal labor scale, we show a robust positive size gradient in tax inspection probability and compliance rate. These gradients, which capture the average effect, mask non-linearities in firm size. Size-based inspection is concentrated among large firms, but appears more stringent for small firms than medium ones, producing a U-shaped relation over the size distribution. This U-shaped pattern is mirrored in the compliance rate. In contrast, the size gradient in informal payments (bribes paid to tax officers) is negative, and also non-linear in firm-size; the largest decrease in informal tax payments occurs over the same segment of the size-distribution where the increase in formal tax compliance is largest. The symmetry provides prima facie evidence of substitution between formal and informal tax payments. Finally, we study heterogeneity across countries and find that size gradients in tax inspection are decreasing with countries' per capita GDP.

Our identification is based on the idea that firm size is partially determined by technological factors (Lucas 1969; Kremer 1993). These technological factors pin down firms' optimal scale of operation (Bain 1954, Burnside 1996, Kumar et al. 1999, Basu and Fernald 2016). If firms in an industry share a common optimal scale across countries, then the relative ranking of two industries' scale, in a plausibly undistorted market such as the US, can serve as an instrument for the relative ranking of the same two industries' optimal scale in a developing country's distorted market.[4] We measure scale as the average number of employees of a firm (N) at the level of three digit ISIC sectors.[5], Our identification strategy relies on predicting the size ranking of industries at 3-digit ISIC level in a WBES country from the size ranking in the US of the same ISIC3 industries. Figure 2 graphically represents the first stage regression, plotting the mean WBES industry size rank within

[3]Questions on tax compliance and informal tax payments are asked at the industry level. Since all our specifications are at the industry-country level, we do not require that the reported answers represents the firm's own tax compliance. Instead we only require that the answers accurately represents the industry's average compliance.

[4]This idea was developed by Rajan and Zingales 1998, who studied if industries reliant on external finance grew faster in countries with more developed financial institutions. They instrument an industry's reliance on external finance with the external finance usage of the same sector in the US.

[5]We show in appendix (Table A1) that results are robust to the employee-weighted average firm size computed in Kumar et al. 1999

each US industry size rank. The slope coefficient in the figure gives the conditional expectation of a WBES country's industry size rank, across ranks of US industry size. The graph suggests that this conditional expectation is very stable and is almost perfectly linear. We show that, when repeated across subsamples of countries at different income levels, this conditional expectation remains very stable and linear and that the slope-coefficient remains constant across subsamples. This suggests that the US distribution has the same power to predict firm size ranking in, for example, Ethiopia, Indonesia, and Brazil. It suggests that the following type of statements is true: "If the average motor vehicles manufacturer requires more workers than the average retail firm in the US, then this ranking of industries by size also holds in Ethiopia, Indonesia, and Bangladesh." The near-perfect linearity and constant slope coefficients across development subsamples suggest our rank-rank identifying strategy captures technological differences in need for labor inputs which vary across industries but not across countries.

The US Census industry size distribution is a valid source of exogenous variation under the identifying assumption that US firms choose their size orthogonally to tax enforcement. This assumption might be reasonable in the US, where the Internal Revenue Service has access to comprehensive sources of third party information and might not need to rely on distortionary size-proxies for tax enforcement. In contrast, tax authorities in in developing countries with low fiscal capacity may be constrained to resort to such size-proxies. contrary to tax authorities in developing countries. We provide support for the identifying assumption. First, the estimates are robust when we use the industry ranking based only on a sample of firms that submit comprehensive financial and accounting forms, instead of the Census.[6] For these firms which disclose substantial amounts of information, size-proxies are arguably no longer determinants of tax inspection. Second, since it has has been documented that tax evasion is important for the self-employed in the U.S. (Blumenthal et al. 2001), we only consider firms with more than five employees in the US Census.[7] Third, for the few countries in the WBES with similar GDP as the US (and thus possibly similar tax enforcement capacity), we show that there is no size gradient in tax inspection.[8]

We uncover sizable size gradients. The IV coefficient implies that a 10 percentile increase in

[6]We use Amadeus data for German and British firms. In Germany, the data derives from annual financial statements, directors' reports and management plans which are filed with the business registry. In Great Britain, the data is derived from audited annual reports which are presented to shareholders.

[7]The selection of firms with five or more employees also matches the sampling strategy of the WBES.

[8]Unfortunately we can not test this directly on the US, since it does not have a World Bank Enterprise Survey.

the WBES size-rank increases a firm's probability of tax inspection by 1.17 percentage points (a 2.6 percent increase relative to mean of 61 percent) and the effective tax rate by 0.61 percentage point (1.8 percent increase). For informal tax payments, a 10 percentile increase in the WBES size-rank decreases the informal tax rate by 0.11 percentage points (7 percent decrease), but has no effect on the extensive margin. The IVs provide a causal estimate of the size-gradients under the first stage identifying assumption and under the exclusion restriction assumption that the US ranking of WBES industry size only impacts tax outcomes through the WBES ranking of industry size. We provide support for the exclusion restriction by showing that the predicted WBES size-ranking is not an imperfect proxy for two competing channels that have been hypothesized to drive tax inspection policies, namely capital intensity and reliance on external finance. (Gordon and Li 2009) Our IV estimates are preciesely estimated, in part owing to the strong first stage F-statistic which is on average in excess of 100. We pursue two main control models which allow non-parametrically for the tax inspection function to vary over firm characteristics or over two digit ISIC industries in every country-year. The latter specification exploits variation in firm size between narrow ISIC3 industries such as "Manufacture of rubber products" (category 251) and "Manufacture of plastics products" (category 252) within ISIC2 category 25. We also find similar coefficients in panel models that exploit variation in the relative industry size-rank of a 3 digit ISIC industry over time, within a country.

The average firm size gradients mask important non-linearities over the size distribution. Size-based inspection and compliance are concentrated among large firms, but appears more stringent for small firms than medium ones, producing a U-shaped relation over the size distribution. The similar pattern of results suggests that size-based tax inspection may explain part of the observed size gradient in compliance. To estimate a direct measure of the effective tax rate, we collect corporate and sales tax rates across countries which we multiply by the compliance rate. We use this new measure to study substitution between informal and formal effective tax rates: the informal tax rate decreases at the same segments of the size-distribution where the formal effective tax rate increases. The magnitude of our results suggests that the size gradient in total effective tax rate (the sum of formal and informal tax rates) remains positive. As a novel result, this substitution mechanism implies that the formal tax rate gradient overstates the total gradient, when the latter is the correct metric for size-induced misallocation.[9]

Finally, we focus on heterogeneity across countries' income levels. We find that the magnitude of the size gradient is decreasing over development. Holding constant the elasticity of firm behaviour with respect to enforcement, this implies that size-driven misallocation is largest in low and middle income countries. The coefficient for countries with similar income levels as the United States is zero and statistically different from the lower income groups. The decreasing reliance on size-based tax policies with development is consistent with empirical evidence showing that countries with weaker fiscal capacity need to rely on production-inefficient tax instruments (Best et al. 2015, Bachas and Soto 2016).

1.1 Related Literature

Our results contribute to three distinct literatures. First, an influential literature analyses cross-country income differences through misallocation of factor inputs across firms and sectors (Hsieh and Klenow 2009). Misallocation can arise from policy-induced distortions, such as size-based policies (Guner et al. 2008, Restuccia and Rogerson 2008, Bartelsman et al. 2013). We show empirically that tax enforcement is a size-dependent policy. Several microeconomic studies focus on size-dependent policies, however their results are specific to a country's context and legislation (Gollin 1995, Harju et al. 2016). An exception is Garicano et al. 2016, which structurally estimates the parameters to construct a counterfactual size distribution. We contribute to this literature by providing reduced form identified estimates of the slope a size policy, which and study its heterogeneity, across the firm size distribution and across countries of different income levels.

Second, our results contribute to the literature on tax enforcement and third-party information (Kopczuk and Slemrod 2006, Gordon and Li 2009, de Paula and Scheinkman 2010, Pomeranz 2015, Naritomi 2016, Jensen 2016). In particular, we provide empirical support to theories where firm size is a key determinant of tax compliance (Kleven et al. 2016, Bigio and Zilberman 2011). In Kleven et al. [2016], tax avoidance strategies such as double book-keeping and collusion with employees to hide operations are impossible to sustain for large firms, since a single whistle blower can bring the entire operation down. As a consequence, large firms disclose third-party informa-

[9]This result also implies that estimating revenue effects of reforms based solely on formal tax returns might not capture the full welfare impacts.

tion and comply with their tax obligations. In Kleven et al. [2016], the government enforcement strategy is fixed, and increased tax compliance is driven by firm's size growth. However, our results are also consistent with models where increased tax inspection with size is an optimal government policy, given fiscal capacity constraints (Bigio and Zilberman 2011, Ito and Sallee 2016).

Third, the estimation of the size gradient in informal payments is related to studies on the incidence of corruption and informal taxation (For a review see Olken and Pande 2012). Svenson [2003] also uses cross-industry variation (in policies and regulations) amongst a sample of Ugandan firms in order to study who pays bribes. Our evidence of decreasing informal tax rates with firm size is consistent with Svenson's finding that bargaining power determines the amount of bribes paid. Our findings of a negative size gradient in effective bribe rate implies that such informal payments are regressive across firms, which complements Olken and Singhal [2011] [10]

The paper is structured as follows. Section 2 provides a stylized model to guide the interpretation of the empirical size gradient. Section 3 discusses the data. Section 4.1 presents the identification strategy and empirical specification. Section 5 shows the main results. Section 6 tests the robustness of the results and assumptions. Section 7 concludes.

2 Conceptual framework

2.1 A reduced form estimate of the size gradient in tax inspection

Size-based tax enforcement, denoted $p(s)$ where p is the (unconditional) probability of receiving a tax inspection and s is firm size, will generate misallocation by distorting firms' optimal choices. It may distort a firm's choices of inputs such as employee-size $n(p(s))$ and output $y(p(s))$.[11] The magnitude of misallocation due to size-based enforcement will depend on two terms. First, it will depend on the elasticity of firm choices with respect to enforcement. Recent bunching studies have exploited threshold-based discontinuities in policy intensity to show substiantial behavioral respones to an increase in tax enforcement (Almunia and Lopez-Rodriguez 2016), labor regulations (Garicano et al. 2016), and accounting requirements (Zareh and Peichl 2016). Second, it will

[10] However our findings are conceptually distinct: the informal payments studied in Olken and Singhal 2011 contribute towards financing public goods and services, and as such can be termed informal tax payments; it is unlikely that the informal bribe payments to tax officials represent an informal source of local public financing.

[11] Size-based enforcement may also lead to distortions through changes to optimal evasion $e(p(s))$, under the condition that evasion is a costly, and thus distortionary, action

depend on the tax inspection gradient in size, denoted $\frac{\partial p}{\partial s}$. Per example, if the return from enforcement is higher for larger than for medium sized firms, enforcement may strongly increase between medium and large firms, leading to a positive size-gradient. The extent of misallocation arising from size-based tax enforcement will depend on the product of the 'behavioral elasticity' and the size-gradient. The aim of this paper is to provide an identified estimate of the average size-gradient estimated over the *full* size-distribution, and to uncover non-linearities in the identified gradient at different points in the size-distribution.[12] The bunching studies rely on a local gradient in enforcement around the size-threshold, $\frac{\partial p}{\partial s} \mid s = s_1$. There are at least two reasons to not extrapolate from $\left(\frac{\partial p}{\partial s} \mid s = s_1\right)$ to derive the distribution-wide gradient.

First, firms at different parts of the size-distribution face different size-based policies within the same country. Indeed, most countries implement enforcement for large firms $s \geq s_L$, but also feature size-based enforcement units for medium taxpayers (for which $s_M < s < s_L$) and small taxpayers (for which $s_S < s < s_M$). Associated with these thresholds are discrete changes in tax inspection p, which due to differences across units in administrative capacity, reporting requirements and type of taxes paid, are likely to differ, such that:

$$\left(\frac{\partial p}{\partial s} \mid s = s_L\right) \neq \left(\frac{\partial p}{\partial s} \mid s = s_M\right) \neq \left(\frac{\partial p}{\partial s} \mid s = s_S\right)$$

It is also not a priori obvious how the effective likelihood $p(s)$ would vary within, say, the medium taxpayer enforcement unit, $\left(\frac{\partial p}{\partial s} \mid s_M < s < s_L\right)$. Suppose the tax inspectors in this unit maximize the expected revenue per inspection, which is the product of evaded income and the likelihood of uncovering the evaded income during an inspection. If more comprehensive information trails for larger firms make it more difficult for them to evade (Kleven et al. 2016, Gordon and Li 2009), then their evaded income could be low but the likelihood of uncovering evasion could be high. Depending on the relative strength of these two forces, expected revenue and the likelihood of a tax inspection could be higher or lower for larger firms.

Second, there may be informal tax inspections (to collect bribes) in parallel to the formal tax inspections (Besley Mclaren). These informal tax payments could increase with firm size if the

[12] Our aim is not to identify behavioral responses to size-driven enforcement. However, by comparing OLS and IV-based estimates of the inspection size-gradient, we find very strongly suggestive evidence that firms change their number employees in reaction to size-based enforcement.

expected bribe is a fraction of the (increasing) tax liability. On the other hand, the informal tax inspection may decrease in size if larger firms have more bargaining power and are more likely to seize the courts. For a given firm size s_i, the total inspection likelihood is therefore the sum of the formal and informal policies.

$$p^{Total}(s_i) = p^{Formal}(s_i) + p^{Informal}(s_i)$$

As a function of how these two size gradients vary with size, the total gradient may be larger or smaller than the formal gradient. The WBES survey measures the total number of visits by tax inspectors, explicitly referencing that informal payments may take place during some of these visits. Given this framing, we interpret the reported visit frequency as a measure of both formal and informal enforcement policies. We thus have an empirical measure across different firm size i of the total tax inspection policy, encompassing both formal and informal enforcement

$$p^{Total} = \sum_i p^{Formal}(s_i) + \sum_i p^{Informal}(s_i) \tag{1}$$

Taken together, this discussion suggests that the sign, magnitude and heterogeneity of the slope of this 'total' inspection policy with respect to size is an open empirical question. Our main objective is to exploit exogenous variation in size to provide an identified estimate of this gradient.

2.2 Why use cross-industry variation in the number of employees?

In order to provide an estimate of the size gradient in tax inspection and tax compliance we need to determine the unit of observations and a measure of firm size. To determine the units of observation, we start with the set of theoretical and empirical papers which investigates the determinants of firm size. The literature can be divided into technological factors (Lucas 1978, Kremer 1993, Kumar et al. 1999) and institutional factors (Hopenhayn 1992 Laporta et al. 1997). The latter includes regulatory barriers to entry (Hopenhayn 1992) and impact of law enforcement (Laporta et al. 1997) and is most consistent with cross-country evidence (Davis and Henrekson 1999). Since we are trying to estimate size-gradients in tax inspection for many countries we instead turn to the technological theories. These theories argue that heterogeneity in production functions yield a size-distribution. Size-heterogeneity could arise from managerial talent (Lucas 1978), labor productivity and diminishing returns to supervision (Rosen 1982), human capital and sorting of workers (Kremer 1993). In practice many studies have used industry level variation to test these theories, arguing that these structural parameters are shared by firms within an industry (Bain 1954, Burnside 1996, Rajan and Zingales 1998 and Kumar et al. 1999). Following this literature we use cross-industry variation, hence determining our unit of observation at the industry level.

Three are three commonly used definitions of firm size: number of employees, output and value-added. In this study, we use the number of employees to measure firm size, since it is likely to be the most accurate and transparent measure of firm size we can obtain from the WBES surveys. In addition to this data observability argument, we should note that a series of papers (surveyed in Kumar et al. 1999) have shown that value added per employee is stable across bins of employee-size, suggesting firm-size measures based on value-added and on number of employees will be very similar. Finally, some technological theories based on coordination costs, are stated in terms of number of employees and not value added nor output (Kremer 1993).[13]

3 Data

To provide global evidence on the relation between firm size and tax inspection and compliance, we use the comprehensive firm-level data from the World Bank Enterprise Surveys (WBES). The surveys cover 125,000 firms in 140 countries between 2003 and 2015. A subset of countries have multiple surveys over time, with an average of 2 surveys per country. The World Bank outsources the data collection to third-party agencies with the objective to remove any official affiliation of the surveyors, which may contaminate the responses. The survey agencies initially draw upon the list of registered establishments provided by the government statistical office. This list is sometimes supplemented with the list of firms registered with the chamber of commerce. The random stratified sampling is first done at the industry-level, typically corresponding to the 2-digit ISIC level, then by firm-size, oversampling from larger firms to capture a large enough share of economic activity. Firms with fewer than five employees, government-owned establishments and co-operatives are dropped from the sampling frame. The WBES contractors contact firms from this stratified-random sample and conduct the surveys with *"the person who most often deals with*

[13] Using number of employees as a proxy for firm size also has as a long intellectual tradition (Pashighian 1968), and is a commonly made choice in other empirical tests of technological size-determinants.

banks or government agencies."

We measure size as the average number of employees per firm in a three digit ISIC sector,[14] excluding part-time and temporary workers. We calculate size for all ISIC3-country-year cells in the WBES. The size distribution in the US is drawn from the 2002 Census of Employment and Wages. To be consistent with the sampling in the WBES, we exclude from the Census firms with five employees or less. While the average firm size is simple and transparent, we also use for robustness an alternative measure of industry size suggested by Davis and Henrekson [1999] and Kumar et al. [1999]. These authors use an employee-weighted size measure, which places additional weight on the firms which have a higher proportion of the total sectoral employment.[15]

We construct extensive and intensive margin measures of tax inspection within an ISIC3-country-year cell as, respectively, the share of firms which reported receiving inspection by tax officials in the past 12 months, and the average number of inspections over that period.

To study (formal) tax compliance, we use the answer to the question:

> *"Recognizing the difficulties many enterprises face in fully complying with taxes and regulations, what percentage of sales would you estimate the typical establishment in your area of activity reports for tax purposes?".*

The compliance rate is the answer to this question, and full compliance is defined as the share of firms that report all of their sales. The reference to a "*typical establishment in your area of activity*" is meant to encourage firms to truthfully report – either a reference group's behavior or the firm's own behavior. While we cannot infer whose behavior the firm is precisely referring to, we only require that the firm's reported compliance rate corresponds to its own ISIC 3 industry, which seems like a plausible assumption, and certainly much weaker than assuming that the answer corresponds to the firm's own compliance rate. We also construct the effective tax rate, defined as the product of the compliance rate and the statutory tax rate (more details in Section 5.3). The tax rate corresponds to the sum of the statutory corporate income tax rate and the general sales tax rate (or VAT).[16]

[14] We use version 3.1 of the ISIC international industry classification.

[15] Specifically, we first calculate the average firm size within size bins and second measure the average size for an ISIC3 as the sum of these bin averages, using as weights the proportion of total ISIC3 employment in that bin.

[16] We collected the statutory sales and corporate tax rates in the relevant year of the WBES country-year cell, using the KPMG worldwide tax summaries and the primary data at www.tradingeconomies.com. The latter source was used whenever a country-year cell was not reported in the KPMG data.

Finally, we define informal payments, with the answer to the question:

> *"It is said that establishments like this one are sometimes required to make gifts or informal payments to public officials to "get things done" with regard to customs, taxes, licenses, regulations. On average, what percent of total annual sales do establishments like this one pay in informal payments or gifts to public officials for this purpose?"*.

This provides a direct measure of the informal tax rate (intensive margin), since the informal payments are expressed as a percentage of sales. The extensive margin of informal payments equals one if a firm makes any informal payments.

Since we study outcomes at the industry level, we report summary statistics at the country-year-industry level, where the industry corresponds to the 3-digit ISIC. In the 272 country year surveys, covering 140 countries, there are on average 50 ISIC3 industries represented, with a median at 52, a 25% percentile at 40 and a 75% percentile at 59. The average industry surveys 10 firms. Table 1 displays the number of observations, mean, standard deviation, and quartiles for each of the variables described above. The average number of tax inspections is just under 2 and 62% of firms receive at least one tax inspection visit in the year. The average compliance rate with taxes is 81% and 57% of firm report being fully compliant. The average informal payment corresponds to 1.6% of firms' sales and 26% of firms make such informal payments. Since the survey oversamples from larger firms it is not surprising to see that 22% of firms are exporters. Similarly since the sampling frame is defined at the two digit ISIC level, manufacturing firms (which occupy a disproportionate share of the ISIC 2 categories) represent 58% of the sample of 3-digit industries. Finally, in the WBES the average ISIC3 industry has 69 workers, while the average 2001 US Census has 53 workers.

Two points are worth highlighting regarding sample size and survey weights. First, while the core tax inspection variable is always available, data on tax compliance is only available for earlier surveys.[17] Second, whenever possible we apply survey weights. However survey weights are missing from some of the earlier surveys. The core tax inspection results are drawn from the sample using survey weights. However to preserve sample size, we report results on tax

[17]The question on formal tax compliance was dropped from the harmonized survey after 2007. In surveys post-2007 administered in Angola, Botswana, Congo (Dem. Rep), Ethiopia, Iraq, Mali and Rwanda, we extracted the tax compliance question from the non-harmonized raw data.

compliance and effective tax rates without dropping observations missing survey weights. We show in the appendix that results remain unchanged in the larger sample without survey weights.

4 Identification and Econometric Specifications

4.1 Identification and First-Stage

The goal of this study is to provide an identified estimate of how firm size impacts tax inspection:

$$\text{Tax inspection}_{ict} = \alpha + \beta \cdot \text{Size}_{ict} + \gamma_{ct} + \varepsilon_{ict} \tag{2}$$

The OLS regression of firm size on tax inspection is likely to suffer from reverse causality and omitted variable bias. In particular, higher tax inspection and effective tax rates for larger firms might induce firms to limit their size. To try to obtain an identified estimate of how firm size impacts the likelihood of tax inspection, we instead use an instrumental variable strategy. A valid instrument predicts firm size and only impacts tax inspection through its effect on firm size. As discussed in Section 2.2, a large theoretical and empirical literature suggests that industries have different optimal scales and that there exists a structural technological demand for labor at the industry level. Our identification strategy follows the intuition of Rajan and Zingales [1998]: if we consider the US as a relatively undistorted market, then US firms achieve their optimal unconstrained size, which depends on the structural scale parameter of their industry. This suggests using the average size of firms in an industry in the US to predict the average size of firms in the same industry in a developing country.

The first stage is estimated with the following regression:

$$\text{Rank size}_{ict} = \alpha_0 + \alpha_1 \cdot \text{Rank size}_{i,US01} + \gamma_{ct} + \varepsilon_{ict} \tag{3}$$

Where Rank size$_{ict}$ is the average firm-size rank of ISIC3 industry i, in country c, at time t, and Rank size$_{i,US,01}$ is the rank by firm-size, of industry i, in the US census in 2001. γ_{ct} are country-year fixed effects. The US industry ranking is drawn relative to the set of industries present in a given country-year, and we weight regressions results by the number of observations in a given

ISIC3-country-year cell.[18] The slope-coefficient α_1 measures the increase in the within country size-ranking of an industry, when going from the smallest to the largest US industry. We choose to implement the first stage using a rank-rank specification for two main reasons. A first advantage is that by using an industry's ranking in terms of average workers, the coefficient β only depends on the joint distribution of average size in the WBES country and average size in the US. Unlike the log-log specification, it does not depend on the marginal distributions of WBES and US industry-size. In other words, in a rank-rank specification β is not impacted by the ratio of US to WBES industry-size variances. The rank-rank specification is thus more stable across subsamples of different development levels with widely varying WBES size-variances.[19] A second advantage is that the rank-rank specification appears more able to untangle technological size-differences from non-technological differences. Non-technological drivers such as availability of labor-saving instruments, labor regulations, and legal quality may differ substantially across countries and impact relative firm size, thus a wedge between WBES level-differences in size and US level-differences in size. Such non-technological drivers of size will not impact β so long as they do not overturn the ranking of industries in WBES relative to the US. The stability of the rank-rank slope-coefficient across development levels is supportive evidence that β is not capturing non-technological determinants (including labor-saving instruments and regulations, legal capacity) which are known to greatly vary across development.[20]

Figure 2 graphically shows the first-stage relation. It displays a non-parametric scatterplot of the relationship between industries' size ranks in the WBES countries and the US Census. In each country-year, ISIC3 industries are ranked relative to other industries in the same country-year survey. The WBES and US Census ranks are then grouped into 50 equal sized (two percentile) bins. Figure 2 plots the mean WBES size rank percentile within each 50-quantile US size rank percentile and the best-fit line. We find a steep positive slope and a linear relation between industries' size-ranks in the US Census and their size-ranks in WBES countries. In Figure 3 we show that across groups of income level, the rank-rank coefficients are constant and very significant. This implies

[18] We do this in order to allow for ISIC3-country-year cells measured with greater precision to carry more weight in the estimates. When omitting such weights, the estimated magnitudes do change but the results are qualitatively the same and remain significant.

[19] This is also the main reason why a rank-rank specification is chosen over the log-log specification in recent studies of mobility (see for example Chetty et al. 2014).

[20] We show in table A1 that our results are robust to running the regressions using the average number of workers per industry rather than their ranking by number of workers.

that the US industry-size distribution has the same power to predict average firm size over the full size-distribution, in countries with vastly different income levels, in for example, Ethiopia, Indonesia and Mexico. In robustness analysis in Section 6, we show that this predictive power does not hinge on using the 2001 Census, and remains almost identical when using either an earlier Census (from 1991) or a more recent Census (from 2015). Finally, the first stage remains strong and positive when we restrict our analysis to the manufacturing sector, a common restriction for studies using cross-industry variation.

The WBES-country size-ranking predicted from the US ranking of the WBES industries is an exogenous source of size-variation under the assumption that US firm-size, as reported in the Census, is orthogonal to tax enforcement. Arguably, for large enough firms, the IRS bases its decision directly on third-party reports and not indirectly based on size. In countries such as the US with high fiscal capacity, it is well-documented that evasion is concentrated among the self-employed and family firms with few employees (Blumenthal et al. 2001, Kleven et al. 2011). To further alleviate concerns, we therefore remove all firms with less than 5 employees in the US Census distribution.[21] In Section 6, we also show that our results hold when constructing the exogenous size-ranking based on a sample of firms in the British and German Amadeus datasets. This data derives from annual financial statements filed with the business registry, and from audited annual reports presented to shareholders. The filing requirements for these firms are subject to more both a broader and deeper set of information reporting requirements, leading tax inspection in this sample not to be likely driven by crude size-proxies. Finally, in section 5.4 we show that for the top income group of the WBES (countries with GDP per capita above $21,000)[22] the estimated coefficients of size on tax inspection is zero. These countries were chosen to be comparable to the US in terms of economic development and fiscal capacity. This evidence supports the first stage identifying assumption that in high fiscal capacity countries, firm size is not driven by size-based inspection.

The IV will provide causal estimates of size-gradient under the first stage validity assumption, discussed above, and the exclusion restruction assumption. The exclusion restriction assumes that the US size-ranking of WBES industries only impacts tax inspection in the WBES country through

[21] Note that the selection of firms with 5 or more employees also matches the WBES sampling frame.
[22] The $21,000 cutoff was chosen to corresponds to the 90th percentile of income in the WBES.

the WBES size-ranking. If tax inspection depends on employee-size only indirectly (e.g. by proxying for output or capital) the exclusion restriction holds as long as technological demand for labor does not impact output or capital other than through labor input. However, this assumption fails if the first-stage coefficient does not capture technological differences in employee-size, but is instead an imperfect proxy for capital input demand. To address this issue we construct two separate measures of demand for capital input in the US industry distribution: the demand for external reliance (Rajan and Zingales 1998) and the capital to labor ratio (Gordon and Li 2009). We show in 6 that when controlling for these capital intensity measures, the coefficients on firm employee size remain unchanged, while the coefficient on capital intensity is insignificant. This suggests that technological demand for labor does not impact tax inspection indirectly through its interaction with capital needs.

4.2 Econometric specifications

The discussion in Section 4.1 motivates the following empirical specifications. The reduced-form size gradient is estimated by directly regressing the WBES industries' tax outcomes on the US industries size rank:

$$\text{Tax outcome}_{ict} = \delta_0 + \delta_1 \cdot \text{Rank size}_{i,US,01} + \gamma_{ct} + \varepsilon_{ict} \tag{4}$$

Where Tax outcome$_{ict}$ is one of the tax outcomes discussed in section 3, such as the likelihood of any tax inspection over the past 12 months. Rank size$_{i,US,01}$ is as previously defined, and γ_{ct} are country-year fixed effects. The coefficient δ_1 identifies the reduced-form size gradient.

The IV specification is:

$$\text{Tax outcome}_{ict} = \beta_0 + \beta_1 \cdot \text{Rank size}_{ict} + \gamma_{ct} + \varepsilon_{ict} \tag{5}$$

Where Rank size$_{ict}$ of industry i in the WBES countries is instrumented with Rank size$_{i,US,01}$ of industry i in the US. In practice, we estimate three different set of models.

The first model combine the US industry variation strategy with a comprehensive set of controls. It allows for tax outcomes to differ non-parametrically and interactively in every country, year, and across a set of industry characteristics. We code all industries as belonging to above or below their country-year median age, share of exporters and share of foreign firms. We then create the matrix $(\text{Characteristics})_{ict}$ containing the full set of interactions across characteristics. The model allows for all factors to impact tax outcomes in an interactive way, resulting in 1,937 fixed effects to be estimated:

$$\text{Tax outcome}_{ict} = \beta_0 + \beta_1 \cdot \text{Rank size}_{ict} + (\text{Characteristics})_{ict} \otimes (\text{Year})_t \otimes (\text{Country})_i + \varepsilon_{ict} \quad (6)$$

The second model allows in a similar interactive way for the tax outcome to differ in every country, year and ISIC2 industry. This implies that variation relies on difference in size of narrow 3-digit ISIC industries within a 2-digit ISIC industry. For example, within ISIC category 25 "Manufacture of rubber and plastics products" it exploits variation in firm size of ISIC3 industries ?Manufacture of rubber products? (Category 251) and ?Manufacture of plastics products? (Category 252). This model estimates 6,130 fixed effects:

$$\text{Tax outcome}_{ict} = \beta_0 + \beta_1 \cdot \text{Rank size}_{ict} + (\text{ISIC2})_{ict} \otimes (\text{Year})_t \otimes (\text{Country})_c + \varepsilon_{ict} \quad (7)$$

Note that in practice, some ISIC2 sectors define the ISIC3 sectors, which leads to a drop in sample size. Further, the drop in size is larger for less developed countries where a smaller degree of specialization implies that some ISIC3 sectors are not represented for a given ISIC2. For transparency, we present results from each model side by side.

The third specification exploits the panel structure of the data, which is available for a subset of countries (the average number of surveys per country is 1.9).[23] In the panel model, we add fixed-effects at the 3-digit ISIC level to the two previous models. Identification comes from variation in industries' relative size ranks within a country and across time. For example the panel model mirroring equation 7 is defined as:

$$\text{Tax outcome}_{ict} = \alpha + \beta \cdot \text{Rank size}_{ict} + (\text{ISIC2})_{ict} \otimes (\text{Year})_t \otimes (\text{Country})_c + \text{ISIC3}_i + \varepsilon_{ict} \quad (8)$$

We report the coefficients β_1 of size on tax outcomes from all three sets of specifications for tax inspection and informal payments. Since the question on tax compliance was discontinued after

[23] Note that in the panel regressions we do not use the instrumental variable strategy.

2007, we have very few repeated country surveys. Therefore we only report the coefficients β_1 of size on tax compliance for the first two models.

5 Results

5.1 Reduced form and IV estimates of size gradients

Tax inspection

In this section we implement the econometric specification described in 4.2. Table 2 reports the size gradient in tax enforcement along the extensive margin (any tax inspection, Panel A) and the intensive margin (number of tax inspections over the past 12 months, Panel B). Panel A shows that industry-size is associated with a higher likelihood of tax inspection. Column 3 and 4 show that the reduced form coefficients are significant in both types of (cross-sectional) fixed effect models. In columns 5 and 6, we estimate the corresponding IV-coefficients for each model. The first stage coefficients are strong (F-statistic of 260 and 105, respectively), and the size gradients are precisely estimated. The coefficient in column 5 implies that a 10 percentile increase in exogenous WBES size-rank is associated with a 2.3 percentage point increase in the likelihood of tax inspection (relative to a mean of 61.8 percent). Columns 7 and 8 exploit the panel dimension of the data, for the subset of WBES countries with multiple surveys. The size gradients are now estimated from changes in inspection associated with a switch in the relative ranking of ISIC3 industries across time. The coefficients in the panel regression are large and significant and of similar magnitude as the IV estimates. Panel B, repeats the above regressions for the number of tax inspections in the past 12 months. The coefficients remain significant and positive across specifications. The IV coefficient in column 5 suggests that a 10 percentile point increase in exogenous WBES size-rank is associated with a 0.14 increase in the number of tax inspections (relative to a mean of 1.98). The panel coefficients (columns 7-8) are again very similar to the IV estimates. To gauge the extent to which the size gradient endogenously determines firm-size, we can compare the OLS to the IV coefficient (respectively, comparing columns 1 and 5, and columns 2 and 6). The IV coefficient is larger in three out of four specifications, suggesting that firms might reduce their size in anticipation of an increased tax inspection likelihood.

Tax Compliance

We now investigate whether we also find a positive size gradient in tax compliance behavior, which could, in part, be explained by increasing tax inspection with firm size. While the identification strategy is the same as for tax inspection, the question on tax compliance only appeared in the first waves of the WBES (from 2003 to 2007), which implies a smaller sample size and insufficient repeated country surveys to run panel regressions. Table 3 shows the size gradient on the extensive and intensive margins of tax compliance. We define the extensive margin as the likelihood of reporting full compliance (Panel A), and the intensive margin as the share of sales reported for tax purpose (Panel B). The IV coefficient in column 5 points to sizable effects of firm size on tax compliance: a 10 percentile increase in the WBES size-rank is associated with a 5.2 percentage point increase in the likelihood of full compliance (relative to a mean of 61.8 percent) and a 2.2 percentage point increase in the share of sales reported for tax purpose (relative to a mean of 80.9). The large wedges between the OLS and the IV coefficients suggests that firms depress their reported size in order to reduce tax compliance.

Informal Payments

In Table 4, we estimate the size gradient in informal payments (bribes) on both the extensive and intensive margins. We define the extensive margin as the likelihood of making any informal payment (Panel A) and the intensive margin as informal tax payments as a share of sales (Panel B), which is the informal effective tax rate. Panels (A) shows no effect of size on the likelihood of making any informal payment. However in Panel (B) we find a large effect on the intensive margin. The IV coefficient in column (5) suggests that a 10 percentile increase in exogenous WBES size-rank is associated with a 0.11 percentage point decrease in the informal tax rate (relative to a base of 1.58%). The negative intensive margin gradient suggests that informal tax payments are regressive across firms. This finding is consistent and complementary to Olken and Singhal [2011], who find that informal tax payments are regressive across households.

5.2 Non-linearity in Size Gradients for Tax Inspection and Tax Compliance

In Section 5.1 we showed that tax inspection and tax compliance increase with firm size, while informal payments decrease. In this section we study potential non-linearities in tax outcomes along the firm-size distribution. This exercise is motivated by the idea that different enforcement instruments, targeting firms of various sizes, translate into a non-linear increase in tax inspection with respect to firm size. To investigate non-linearities, we analyze how the reduced form relation between a tax outcome and the industry firm size ranking changes along the size distribution. To this end, we first residualize the industry size-ranking in the US Census, N_{US}, with respect to controls and the country and year fixed effects in Equation 6. Similarly, we residualize the tax outcome of interest (e.g. tax inspection likelihood) with respect to the same controls and fixed effects. We split the residualized industry size-ranking $\widehat{N_{US}}$ into twenty groups by size (vingtiles). Figure 4 shows the median of the residualized tax outcomes for each vingtile of industry-size. These binned scatterplot provide a non-parametric representation of the estimated reduced-form coefficient.

Panels A and B of Figure 4 plot the average size gradient, respectively, for the likelihood of tax inspection and for the tax compliance rate. In Panel A, the tax inspection likelihood decreases with size at the very bottom of the distribution, remains flat around the middle part, and steeply increases in the top third of the distribution. In Panel B, the formal compliance rate is also U-shaped over the size distribution and closely mirrors the changes in tax inspection likelihood, especially in the top third of the size distribution. This suggests that the positive size gradient in formal compliance is partly driven by size-dependent tax inspection.

5.3 Substitution Between Formal and Informal Tax Rates

The estimation of the coefficients on tax compliance and informal payments yielded opposite signs. This is suggestive of substitution behavior between formal and informal taxation. In order to directly test for this substitution behavior, we need a measure of firms' informal and formal effective tax rates. Since firms report informal payments as a share of their sales, we directly have the effective informal tax rate. To obtain the effective formal tax rate, we take the product of the compliance rate with the country's statutory tax rate. We collect corporate and sales tax rates for all the country-year pairs where this information was available. We then construct the effective formal tax rate of an ISIC-2 industry by multiplying its compliance rate with the sum of the corporate and sales tax rates. We recognize several limitations with this methodology. First firms of different size could face different statutory tax rates. While this is rare for the sales taxes, the corporate income tax sometimes grants lower rates to small and medium firms. Second we are assuming that the answer to the question on tax compliance applies to both the sales and corporate income tax, while the question does not refer to any specific tax. Third, we do not collect information on other taxes (e.g. payroll taxes) which could be relevant in practice. With these caveats in mind, we view this section as a tentative to study substitution patterns across formal and informal taxes and treat the results with caution.

In Panel A of Figure 5 we plot the residualized formal and informal effective tax rates, for each vingtile of the firm size distribution. There appears to be a substitution between formal and informal effective tax rates: the informal tax rate rises (falls) over the portions of the size-distribution where the formal tax rate falls (rises). This result is an equilibrium outcome and could be driven by a combination of demand and supply side factors. For example, holding tax inspector behavior constant, this substitution pattern could occur if larger firms are more constrained in evading formal taxes while having higher bargaining power in informal tax negotiations. Panel A shows that, as a percent of total sales, the formal tax payments are much larger in absolute terms than the informal payments. However, when dividing the residualized rates by the respective average rates, Panel B reveals that firm size explains a much larger share of the variation in informal taxes than for formal taxes.

Table 5 displays the size gradients for the formal effective tax rate (Panel A) and the total effective tax rate (Panel B), which equals the sum of the formal and informal rates. Panel A shows that the average size gradient for the formal tax rate is positive and significant. The IV-coefficient in column 5 implies that a 10 percentile rank increase leads to a 1.1 percentage point increase in the firms formal effective tax rate (relative to a mean of 31%). The size gradient in the total effective tax rate is positive, but not statistically significant. Its magnitude is consistent with the sum of the (negative) informal gradient and the (positive) formal gradient. The statistical insignificance of the total effective rate could be due to our imperfect proxy for the statutory tax rate. These results support the hypothesis of substitution between formal and informal taxation and highlight the importance of informal payments for studies measuring tax evasion in developing countries.

5.4 Heterogeneity across development levels

Our final empirical exercise investigate differences in the size gradient over levels of economic development. Reliance on size-dependent policies for tax enforcement could originate in state capacity: at lower levels of development the tax authority is constrained by a lack of information on firms and has to resort to imperfect proxies such as size-dependent polices (Best et al. 2015). As a country's fiscal capacity grows (Kleven et al. 2016, Jensen 2016), the tax authority's reliance on size-dependent policies can be weakened. We therefore hypothesize that the size gradient in tax inspection decreases with a country's income level. To test this hypothesis, we estimate the tax inspection size gradient across subsamples of countries at different levels of development.[24] Figure 6 presents the results of these regressions: Panel A shows the reduced form and IV coefficients for the likelihood of tax inspection at each income level, while Panel B shows the first stage. In Panel A the magnitude of the tax inspection size gradient is decreasing over levels of development, both in the reduced form and IV specifications. The IV coefficients are large and significant for lower-middle to upper-middle income levels.[25] These levels of development include the countries often studied in the misallocation literature (e.g. India, Mexico, China) and we show that for these countries tax inspection is a policy-induced size distortion. On the contrary, in high-income countries both the reduced form and IV coefficients are centered at zero. This is despite the first stage coefficient being constant across income levels (Panel B) and the 1st stage F-statistic being well above 10. The absence of size-dependent policies at high income levels is consistent with the theory that countries with strong fiscal capacity are able to decrease their reliance on production-inefficient enforcement instruments. It also provides some support to the identifying assumption that tax inspection is orthogonal to size in a high tax capacity environment.

[24] Country's income groups follow the World Bank economic classification, with the exception of the top group, chosen as the 90th percentile of per capita GDP distribution of countries. This cutoff was decided in order to have in the top group countries most similar to the US.

[25] We should note that the coefficients for the lowest income countries (GDP per capita below $1,100) are not significant at the 5%, although the IV coefficient is significant at the 10% level. There are two explanations for this. First, this group only contains fewer country-year observations than the lower-middle, middle and higher-middle income groups. Second we are weighting ISIC-3 industries by their number of observations. Surveys in the poorest countries tend to be smaller in terms of the number of firms and appear to have less economic sectors represented.

6 Robustness Analysis

6.1 Robustness and validity of instrumental variable

In this section we show that our main results on tax inspection are robust to the choice of the instrumental variable. In the core results we instruments the average firm-size rank at the ISIC3 level using the average firm-size rank in the 2001 US Census. Table A1 presents the first stage and IV coefficients on tax inspection, for six different definitions of the instrumental variable. Columns 1-3 show results using each of the last three waves of the US firm Census (1991, 2001, 2015). The Census size-rank across the three samples provides the same linear predicted ranking of WBES size and similar IV coefficients. We also show that the first stage is constant at different levels of development for each census wave (Figures A1 & A2). Since the validity of the instrument rests on the assumption that the firm-size distribution is undistorted by tax enforcement, using industry ranking from other high fiscal capacity OECD countries should yield similar results. Furthermore, it is likely that size proxies for tax enforcement are used by tax authority among firms with stringent reporting requirements. Therefore, we also construct the 1st stage size-ranking based on a set firms in Germany and the UK using the Amadeus database. In Germany, the data derives from annual financial statements, directors' reports and management plans which are filed with the business registry. In Great Britain the data is derived from audited annual reports which are presented to shareholders. These firms have submitted comprehensive balance sheet and profit-loss information. Column (4) in Table A1 shows that the Amadeus size-ranking predicts a similar linear WBES size-ranking, with a constant slope at different levels of development (Figure A3). Finally Columns 5 and 6 of Table A1, use alternative definitions of firm-size in the US Census 2001. Column 5, uses size-ranks that are employee-weighted to give more weight to firms which concentrate large shares of total employment (Kumar et al. 1999). Column 6, uses the average firm-size per per industry, instead of the industry-size ranking. The 1st stage rank-rank is similar across alternative samples, despite the well-documented differences in size-determinants across US and European firms, and across US firms over time. This suggests that our identifying variation satisfies the exclusion restriction and is identified from country-invariant technological differences in labor input. Finding the same magnitude in the Amadeus sample, where firms submit comprehensive financial and accounting forms and are subject to pervasive information trails regardless of size, supports the validity of our instrumental variable.

6.2 Additional controls and samples

Is worker-size proxying for capital intensity?

A potential limitation of our identification strategy is that firm size, as measured by its number of workers, could proxy for firm's capital intensity. Therefore what we capture is not the direct impact of firm size, but instead the impact of increased capital on tax inspection. To mitigate this concern, we construct an industry capital intensity ranking, using (1) the measure of reliance on external funds Rajan and Zingales [1998] and (2) the capital to labor ratio as in Gordon and Li [2009]. We then regress jointly the firm size ranking and capital ranking measures on tax inspection (Table A3). The inclusion of controls for capital intensity does not impact the results, and the coefficients on capital rankings are very small in magnitude and insignificant. This suggests that the size rank distribution captures technological demand for labor and not demand for capital. As an additional robustness check, we use the reduced form equation to show that that the US-size ranking is associated with outcomes that are predicted by theory to vary with demand for labor input (Lucas 1969; Kremer 1993). In particular, we find that an increase in US Census size-rank predicts higher sales growth, higher labor cost per permanent employee, higher perceived constraints due to labor regulation, increased likelihood of having internationally recognized quality certification and higher likelihood of being part of a larger firm.

Results using only the manufacturing sector

We try to exclude the possibility that results are driven by a small set of peculiar ISIC3 industries. To do so, we limit the sample to the manufacturing sector, which is over-sampled in the WBES, and where most ISIC2 codes are well represented across countries and (Manufacturing represents 60% of the WBES firms). Table A2 shows the reduce form and IV results on the intensive and extensive margins of tax inspection. The 1st coefficient is exactly the same as the coefficient in the full sample (0.310 in manufacturing versus 0.308 in the full sample). The reduced form coefficients are larger than in the full sample despite a similar-sized first stage, which could suggest that the size gradient within manufacturing is larger than in the remaining sectors. In any case, the overall size and significance of coefficients are in line with the results from the full sample, alleviating concerns that these results were driven by a small set of outlier ISIC3 industries.

6.3 Size gradient interactions with statutory tax rates

When the statutory tax rate is high, firms might evade more taxes and pay larger informal bribes. For the tax authority this implies that the returns to size based taxation might be larger, and the gradients in tax inspection larger. To test this hypothesis we regress tax outcomes on the interaction between rank-size and a high tax dummy variable. The high tax dummy is equal to one if the country's statutory tax rates is above the sample's median, where we defined the statutory tax rate as the sum of the corporate and sales tax rates. Table A4 shows that the interaction term between the industry size rank and the high tax dummy is positive for tax inspection and compliance and negative for informal payments. This suggests that in countries with high tax rates, the tax inspection and tax compliance gradients are steeper, while the substitution with informal taxation is stronger.

7 Conclusion

In this paper we study how firm size impacts tax inspection and tax compliance across all levels of development. The identification strategy uses the relative ranking of industries' average firm-size in the US as an instrument for the relative ranking of the same industries' size in a developing country. We uncover a sizable positive size gradient in both tax inspection and tax compliance. This result is very robust to using high-dimensional fixed-effects allowing tax inspection functions to vary flexibly across firm characteristic-country-year cells, variation within narrowly defined industries, and variation within industries over time. We show that while compliance rates increases with size, informal payments decrease with size, generating a substitution pattern between formal and informal tax payments. We also document that the relation between firm size and tax outcomes is non-linear. In particular, the size-based inspection policies and compliance behavior are particularly concentrated among large firms and more stringent for small firms than medium firms. These results provide a robust empirical foundation of a systematic size-based policy which can generate wedges across firms of different sizes. We view our results as complementary to a growing set of within-country studies that exploit local variation in enforcement around a size-based threshold to estimate behavioral elasticities. Our empirical exercise allows us to identify the enforcement size-gradient across the size distribution and across development levels. However, it is not informative about the magnitude of the behavioral elasticities associated with the gradients. The local elasticity must be multiplied by the local gradient at all parts of the size distribution to generate an aggregate estimate of size-driven distortions; this could be pursued in future work to generate estimates of misallocation across development. Our results provide empirical support to models which predict a positive relation between firm size and tax inspection or compliance (Kleven et al. 2016, Bigio and Zilberman 2011). One normative prediction that arises from these models is that non-linear tax inspection along the firm-size distribution may arise as an optimal policy, in spite of the potential size distortions it generates. This paper does not participate in the normative debate, but might help in understanding tax policy in low fiscal capacity environments, where both positive and normative issues remain active areas of research.

References

Almunia, Miguel and David Lopez-Rodriguez, "Heterogeneous Responses to Effective Tax Enforcement: Evidence from Spanish Firms," 2016.

Bachas, Pierre and Mauricio Soto, "Not(ch) Your Average Tax System: Corporate Taxation Under Weak Enforcement," 2016.

Bain, Joe S., "Economies of Scale, Concentration, and the Condition of Entry in Twenty Manufacturing Industries," *American Economic Review*, 1954, *44* (1), 15–39.

Bartelsman, By Eric, John Haltiwanger, and Stefano Scarpetta, "Cross-Country Differences in Productivity : The Role of Allocation and Selection," *American Economic Review*, 2013, *103* (1), 305–334.

Basu, Susanto and John G Fernald, "Returns to Scale in U.S. Production: Estimates and Implications," *Journal of Political Economy*, 2016, *105* (2), 249–283.

Best, Michael, Anne Brockmeyer, Henrik Kleven, Johannes Spinnewijn, and Mazhar Waseem, "Production vs Revenue Efficiency With Limited Tax Capacity: Theory and Evidence From Pakistan," *Journal of Political Economy*, 2015, *123(6)* (9717), 1311–1355.

Bigio, Saki and Eduardo Zilberman, "Optimal Business Income Tax Enforcement," *Journal of Public Economics*, 2011, (June), 1021–1035.

Blumenthal, Marsha, Charles Christian, and Joel Slemrod, "Do Normative Appeals Affect Tax Compliance? Evidence from a Controlled Experiment in Minnesota," *National Tax Journal*, 2001, *54*, 125–38.

Burnside, Craig, "Production function regressions, returns to scale and externalities," *Journal of Monetary Economics*, 1996, *37*, 177–201.

Chetty, Ray, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez, "Where is the Land of Opportunity? the Geography of Intergnerational Mobility in the United States," *Quarterly Journal of Economics*, 2014.

Davis, Steven J and Magnus Henrekson, "Explaining National Differences in the Size and Industry Distribution of Employment," *Small Business Economics*, 1999, *12* (1), 59–83.

de Paula, Aureo and Jose A Scheinkman, "Value Added Taxes, Chain Effects and Informality," *American Economic Journal: Macroeconomics*, 2010, *2* (October), 195–221.

Garicano, Luis, Claire LeLarge, and John Van Reenen, "Firm Size Distortions and the Productivity Distribution: Evidence from France," *American Economic Review*, 2016, *106* (18841), 3439–3479.

Gollin, Douglas, "Do Taxes on Large Firms Impede Growth? Evidence from Ghana," *Bulletins 7488, University of Minnesota, Economic Development Center.*, 1995.

Gordon, Roger and Wei Li, "Tax structures in developing countries: Many puzzles and a possible explanation," *Journal of Public Economics*, 2009, *93* (7-8), 855–866.

Guner, Nezih, Gustavo Ventura, and Yi Xu, "Macroeconomic implications of size-dependent policies," *Review of Economic Dynamics*, 2008, *11* (4), 721–744.

Harju, Jarkko, Tuomas Matikka, and Timo Rauhanen, "The effects of size-based regulation on small firms: evidence from VAT threshold," *CESifo Working Paper No 6115*, 2016.

Hopenhayn, Hugo, "Entry , Exit , and firm Dynamics in Long Run Equilibrium," *Econometrica*, 1992, *60* (5), 1127–1150.

Hsieh, Chang-Tai and PJ Klenow, "Misallocation and Manufacturing TFP in China and India," *Quarterly Journal of Economics*, 2009, *CXXIV* (November).

Ito, Koichiro and James M Sallee, "The Economics of Attribute-Based Regulation: Theory and Evidence from Fuel-Economy Standards," 2016.

Jensen, Anders, "Employment Structure and the Rise of the Modern Tax System," 2016.

Kanbur, Ravi and Michael Keen, "Thresholds, informality, and partitions of compliance," *International Tax and Public Finance*, 2014, *21* (4), 536–559.

Kleven, Henrik Jacobsen, Claus Thustrup Kreiner, and Emmanuel Saez, "Why can modern governments tax so much? an agency model of firms as fiscal intermediaries," *Economica*, 2016, (15218), 51.

_ , Martin B Knudsen, Claus Thustrup Kreiner, S\oren S\oren Pedersen, and Emmanuel Saez, "Unwilling or Unable to Cheat? Evidence From a Tax Audit Experiment in Denmark," *Econometrica*, 2011, *79* (3), 651–692.

Kopczuk, Wojciech and Joel Slemrod, "Putting firms into optimal tax theory," *American Economic Review Papers and Proceedings*, 2006, *96* (2), 130–134.

Kremer, M., "The O-Ring Theory of Economic Development," *The Quarterly Journal of Economics*, 1993, *108* (3), 551–575.

krishna Kumar, Raghuram Rajan, and Luigi Zingales, "What Determines Firm Size?," *NBER WORKING PAPER SERIES*, 1999, *Working Paper 7208.*

Laporta, Rafael, Florencio Lopez de silanes, Andrei Shleifer, and Robert W Vishny, "Legal Determinants of External Finance," 1997.

Lucas, Robert, "Labor-capital substitution in US manufacturing.," in "In Harberger, A. C. and Bailey, M. J., editors, The taxation of income from capital. Washington, D. C.: Brookings Institution." 1969.

Lucas, Robert E., "On the Size Distribution of Business Firms," *The Bell Journal of Economics*, 1978, *9* (2), 508–523.

Naritomi, Joana, "Consumers as Tax Auditors," 2016.

Olken, Benjamin A and Monica Singhal, "Informal Taxation," *American Economic Journal: Applied Economics*, 2011, *3* (October), 1–28.

Olken, Benjamin A. and Rohini Pande, "Corruption in Developing Countries," *Annual Review of Economics*, 2012, *4* (1), 479–509.

Pashighian, B, "Limit Price and the Market Share of the Leading Firm," *Journal of Industrial Economics*, 1968, *16* (3).

Pomeranz, Dina, "No Taxation without Information: Deterrence and Self-Enforcement in the Value Added Tax," *American Economic Review*, 2015, (8), 2539–2569.

Rajan, Raghuram G and Luigi Zingales, "American Economic Association," *The American Economic Review*, 1998, *88* (3), 559–586.

Restuccia, Diego and Richard Rogerson, "Policy distortions and aggregate productivity with heterogeneous establishments," *Review of Economic Dynamics*, 2008, *11* (4), 707–720.

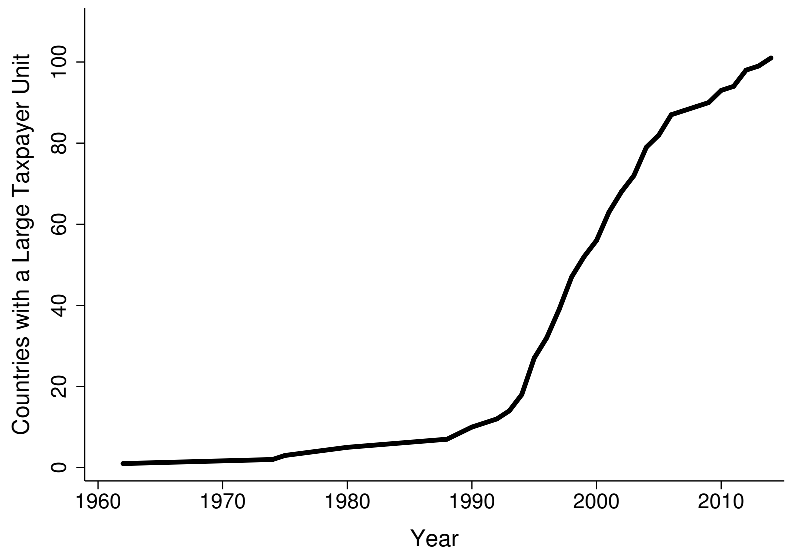
Rosen, Sherwin, "No Title," *Bell Journal of Economics*, 1982, *13* (2), 311–323.

Svenson, Jakob, "Who Must Pay Bribes and How Much ? Evidence from a Cross Section of Firms," *Quarterly Journal of Economics*, 2003, *118* (1), 207–230.

World Bank, "Doing Business 2017: Equal Opportunity for All.," *Washington, DC: World Bank. DOI: 10.1596/978-1-4648-0948-4. License: Creative Commons Attribution CC BY 3.0 IGO*, 2017.

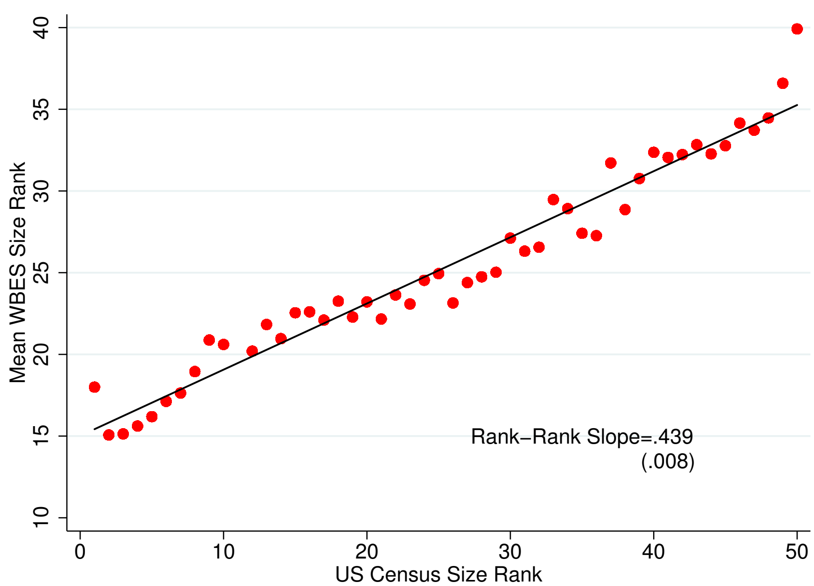
Zareh, Asatryan and Andreas Peichl, "Responses of Firms to Tax, Administrative and Accounting Rules: Evidence from Armenia," 2016.

FIGURE 1: NUMBER OF COUNTRIES WITH A LARGE TAXPAYER UNIT OVER TIME



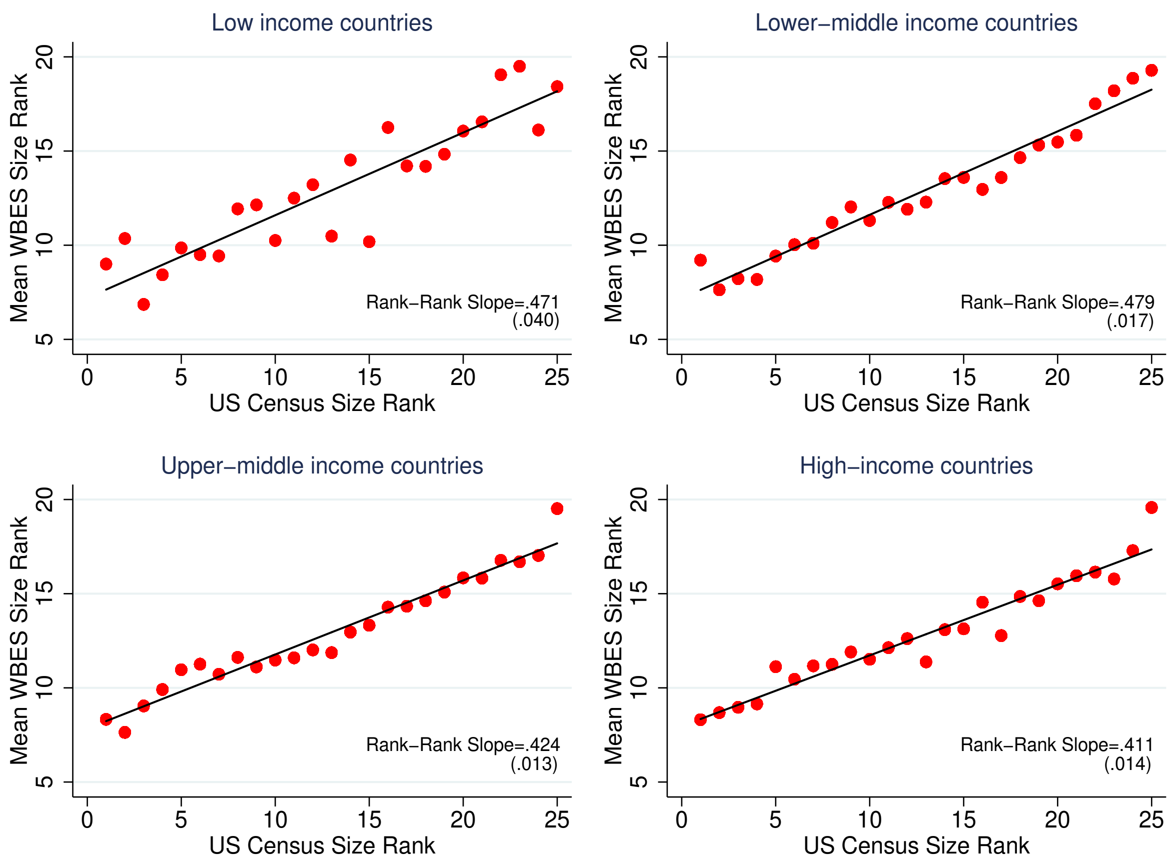
Source: Data collected by the authors. Sample of 113 countries, with more than one million inhabitants.
Figure 1 shows that in the past twenty years more than a 100 countries adopted a Large Taxpayer Unit (LTU). This development is part of an increased trend of taxpayer segmentation, recommended by international institutions.

FIGURE 2: INDUSTRY SIZE IN THE US ON INDUSTRY SIZE IN THE WBES [FIRST STAGE]



Source: US Census 2001 and World Bank Enterprise Surveys 2003-2015.
Figure 2 presents non-parametric binned scatterplot of the relationship between WBES countries and US firm size ranks, as discussed in Section 4.1. WBES size rank is the rank of an ISIC3 by number of permanent employees per firm in the WBES country-year. US Census size rank is the ISIC3 rank of number of employees per firm in 2001, excluding firms with less than 5 employees. In each country-year, ISIC3 industries are ranked relative to other industries in the same country-year survey. The WBES and US Census ranks are then grouped into 50 equal sized (two percentile) bins. This graph plots the mean WBES size rank percentile within each 50-quantile US size rank percentile and the best-fit line.

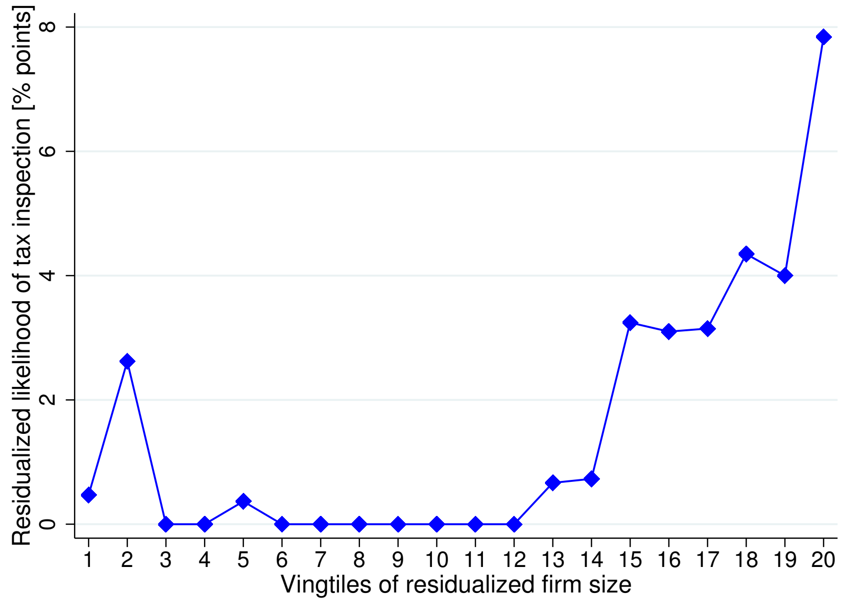
FIGURE 3: INDUSTRY SIZE IN THE US ON INDUSTRY SIZE IN THE WBES BY INCOME LEVEL



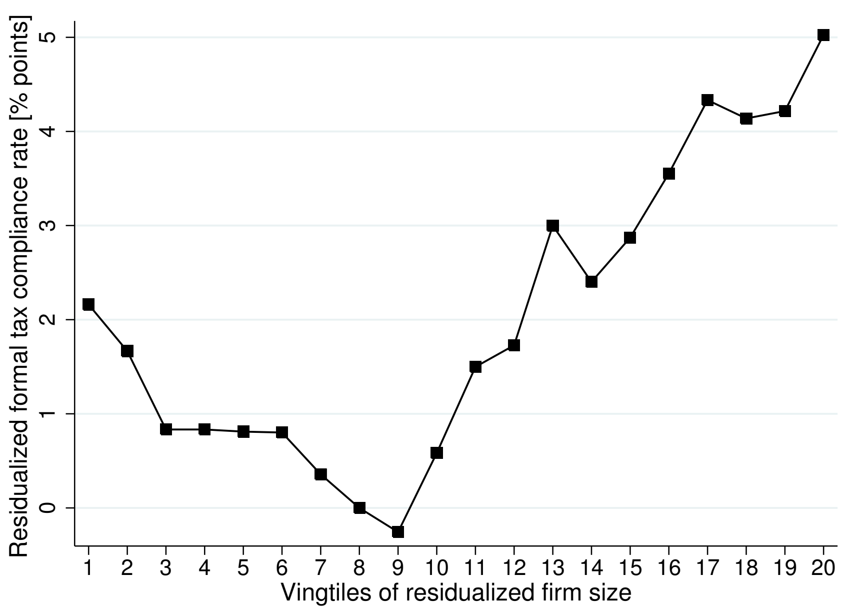
Source: US Census 2001 and World Bank Enterprise Surveys 2003-2015.
Figure 3 presents non-parametric binned scatterplot of the relationship between WBES countries and US firm size ranks, in the four per capita income groups based on the World Bank income classification. These results are discussed in Section 4.1 The cut-off per-capita income values are: 1100 for low income countries; between 1100 and 3900 for lower-middle income countries; between 3900 and 12730 for upper-middle income countries; above 12730 for high income countries. WBES size rank is the ISIC3 rank of number of permanent employees per firm in the WBES country-year. US Census size rank is the ISIC3 rank of number of employees per establishment in 2001, where we exclude establishments with less than 5 employees. In each country-year, ISIC3 industries are ranked relative to other industries in the same country-year survey. Within each per capita income group, the WBES and US Census ranks are then grouped into 25 equal sized bins. For each per capita income group, this graph plots the mean WBES size rank percentile within each 25-quantile US size rank percentile, together with a best-fit line. The rank-rank slope coefficient and standard error are estimated in each per capita income group.

FIGURE 4: NON-LINEAR IMPACT OF FIRM SIZE ON TAX INSPECTION & TAX COMPLIANCE

(A) PANEL A: TAX INSPECTION



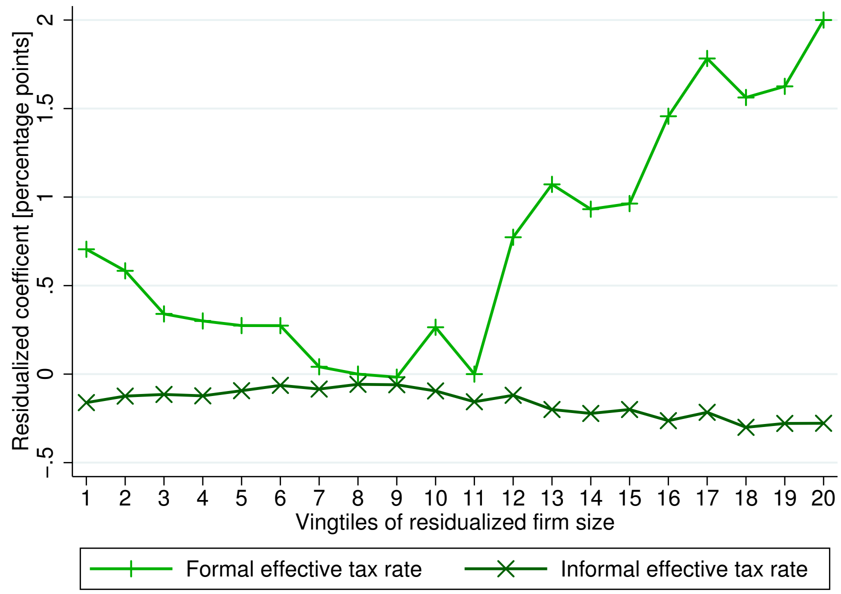
(B) PANEL B: TAX COMPLIANCE



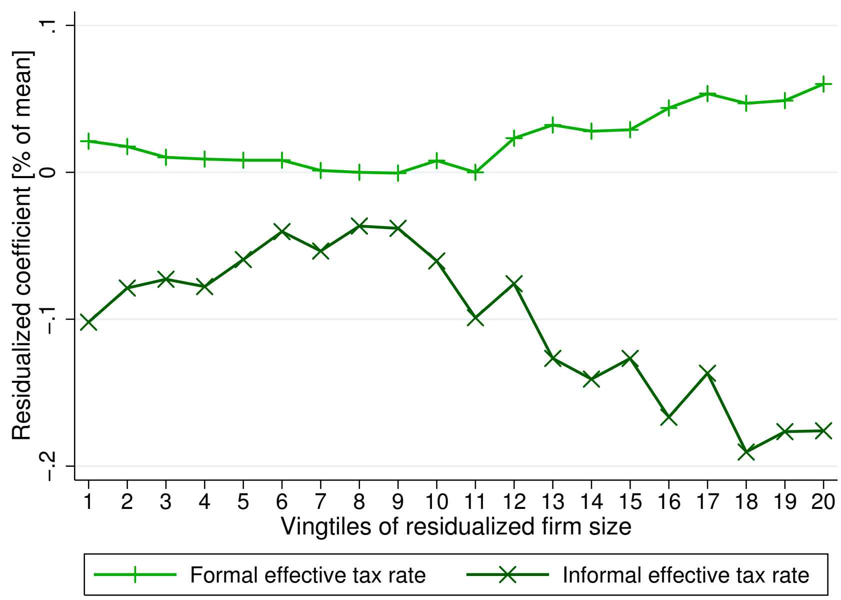
Source: US Census 2001 and World Bank Enterprise Surveys 2003-2015.
Figure 4 presents binned scatter plots of the likelihood of tax inspection and formal tax compliance rate versus firm size, discussed in section 5.2. To construct these panels, we first residualize the y-axis and x-axis variables with respect to the country-year-control interactive fixed effects. We then group the residualized firm-size into twenty equal-sized bins (vingtiles) and plot the median of the y-variable residual within each bin. The median is chosen to reduce the impact of residual outliers. The specification corresponds to the regressions in Column 3 of Table 2 (Panel A) and Column 3 of Table 3 (Panel B) respectively, and use the same sample restrictions and variable definitions.

Figure 5: Non-Linear Impact of Firm Size on Formal & Informal Tax Rates

(a) Panel A



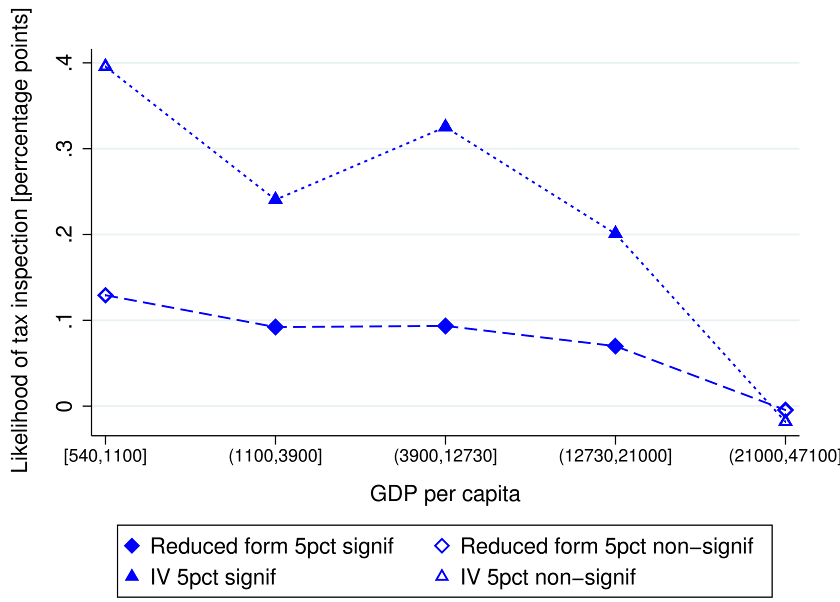
(b) Panel B: as Percentage of Mean of Outcomes



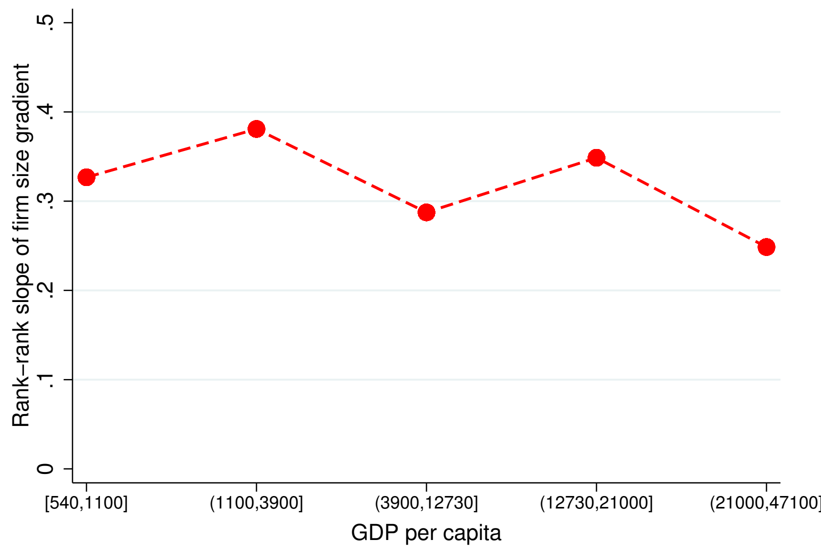
Source: US Census 2001 and World Bank Enterprise Surveys 2003-2015.
Figure 5 shows binned scatter plots of the effective formal and informal tax rates on firm size, discussed in section 5.3. These plots correspond to the regressions in Column 3 of Table 3 Panel B and Table 4 Panel A respectively, and use the same sample restrictions and variable definitions. To construct these panels, we first residualize the y-axis and x-axis variables with respect to the country-year-control interactive fixed effects (defined in Section X equation Y). We then group the residualized firm-size into twenty equal-sized bins (vingtiles) and plot the median of the y-variable residual within each bin. The median is chosen to reduce the impact of residual outliers.

FIGURE 6: SIZE GRADIENTS AND INDUSTRY SIZE RANK ASSOCIATIONS BY INCOME LEVELS

(A) PANEL A: TAX INSPECTION



(B) PANEL B: SIZE RANK-RANK ASSOCIATION [1ST STAGE]



Source: US Census 2001 and World Bank Enterprise Surveys 2003-2015.
Figure 6 present scatterplots of the rank-rank size relationship and the size-tax inspection relationship in different per capita income groups, discussed in Section 5.4. The groups correspond to the World Bank income classification, except for the top income group (above USD 21,000) which corresponds to the 90th percentile of income in the WBES and was chosen such that countries are comparable to the US in terms of economic development. In Panel A, we show the reduced form and IV coefficients, replicating the regressions in Columns 3 and 5 of Table 2 for each development level. In Panel B, we report the coefficients from the rank-rank size regression (Firs-stage) for each of the above IV specifications. The diamond series corresponds to the reduced form regression, while the triangle series corresponds to the IV regression. Filled symbols indicate statistical significance at the 5% level, while closed symbols indicate non-significant results at the 5% level.

TABLE 1: SUMMARY STATISTICS

Variable	(1) N	(2) Mean	(3) s.d.	(4) p25	(5) p50	(6) p75
Panel A: Industry Level Outcomes (ISIC3)						
Any Tax Inspection [0,100%]	12,152	61.88	35.65	37.50	66.67	100
Number of Tax Inspections	12,152	1.98	2.13	0.67	1.48	2.72
Full Compliance [0,100%]	3,343	57	38.1	25	60	100
Compliance Rate	3,343	80.93	23.6	70	89.6	100
Any Informal Payments [0,100%]	10,013	26.00	36.41	0	0	50
Informal Payments, Share of Sales	10,013	1.58	4.22	0	0	1.50
Formal Tax Payments, Share of Sales	3,048	33.28	11.9	25.8	35	42
Formal and Informal Tax Payments, Share of Sales	2,864	34.97	12.04	27.63	36	42.46
Panel B: Inudustry Level Covariates (ISIC3)						
Age	12,152	18.64	13.16	10.75	15.57	22.67
Share of Establishment Owned by Foreign Companies	12,149	0.137	0.266	0	0	0.15
Share of Annual Sales Destined for Direct Export	12,093	0.228	0.326	0	0	0.36
Manufacturing Sector {0,1}	12,152	0.58	0.49	0	1	1
Statutory Corporate Tax Rate + Sales Tax Rate	11,524	0.39	0.09	0.35	0.40	0.45
GDP Per Capita, Constant USD 2011	12,117	10,673	8,554	3,736	8,837	16,459
Panel C: Industry Size Characteristics (ISIC3)						
WBES Size	12,146	69.42	289.5	12.94	24.66	52.84
US Census 1991 Size	11,330	48.2	39.09	27.33	38.28	66.18
US Census 2001 Size	11,962	52.95	43.73	28.20	40.48	67.95
US Census 2015 Size	11,958	51.65	40.89	27.79	37.40	62.47
EU Amadeus 2007 Size	12,032	393.7	891.7	116.76	184.05	346.84
WBES RKZ-Weighted Size	12,146	262.3	840.3	21.06	67.40	200.57
US Census 2001 RKZ-Weighted Size	12,001	302.2	306.60	125.58	218.41	384.67

Source: World Bank Enterprise Surveys 2003-2015.
Table 1 displays the summary statistics at the industry-country-year level, as described in section 3. Industries are defined at the 3-digit ISIC level, using the UN classification version 3.1. Column 1 reports the number of industry level observations, column 2 the mean, column 3 the standard deviation and columns 4-6 the quartiles (25% percentile, median and 75% percentile). All statistics are reported for the full sample which includes some observations for which a sampling weights are missing.

TABLE 2: IMPACT OF FIRM SIZE ON TAX INSPECTION

	OLS		Reduced Form		IV		Panel: OLS	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A								
Outcome: Tax Inspection {0,1}								
Industry Size Rank (World Bank)	0.197***	0.232***			0.234***	0.189*	0.303***	0.276***
	(0.054)	(0.043)			(0.055)	(0.109)	(0.065)	(0.067)
Industry Size Rank (US Census)			0.073***	0.050*				
			(0.017)	(0.030)				
FE country*year*controls	✗		✗		✗		✗	
FE country*year*ISIC2		✗		✗		✗		✗
FE ISIC3							✗	✗
F-Stat					259.68	105.49		
R-squared	0.650	0.772	0.641	0.766	0.650	0.772	0.890	0.906
Panel B								
Outcome: # of Tax Inspections								
Industry Size Rank (World Bank)	0.0112***	0.0161***			0.0142***	0.0181***	0.0170***	0.0155***
	(0.0030)	(0.0025)			(0.0034)	(0.0056)	(0.0046)	(0.0038)
Industry Size Rank (US Census)			0.0045***	0.0048***				
			(0.0010)	(0.0017)				
			(0.017)	(0.030)				
FE country*year*controls	✗		✗		✗		✗	
FE country*year*ISIC2		✗		✗		✗		✗
FE ISIC3							✗	✗
F-stat					259.68	105.49		
R-squared	0.600	0.734	0.591	0.725	0.599	0.734	0.886	0.896
Observations	9,772	7,864	9,772	7,864	9,772	7,864	5,048	4,176
Number of clusters	131	131	131	131	131	131	75	77

Source: US Census 2001 and World Bank Enterprise Surveys 2003-2015.
Robust standard errors in parentheses, clustered at the country level. *** p<0.01, ** p<0.05, * p<0.1
Table 2 shows the coefficients of regressions of industry firm-size ranking on tax inspection, discussed in section 5.1. Panel A measures tax inspection as a dummy which takes the value one if the firm was visited during the year. Panel B measures tax inspection as the number of times the firm was visited by tax inspectors. World Bank size rank is the ISIC 3-digit rank of number of permanent employees per firm in the WBES country-year. US Census size rank is the ISIC 3-digit rank of number of employees per establishment in 2001 quarter 1, where we exclude establishments with less than 5 employees. Odd number columns include interactive fixed effects between country, year and 3 dummies indicating whether the ISIC3 industry was above median in the country-year distributions of age, export-share, and foreign-industry share. Even number columns include interactive fixed effects between country, year and 2-digit ISIC codes. In this case, size-rank coefficients are estimated using variation between 3-digit ISIC codes, controlling non-parametrically for country-year-2digit ISIC effects. For a given specification, the sample size in the even columns is smaller than in the odd columns. This is because the ISIC2 interactive fixed effect specification drops all cells where there does not exist more than one ISIC3-country-year observation within a ISIC2-country-year observation. The sample drop occurs in industries that are less specialized, which are more prevalent in less developed countries. In columns (5) & (6), we instrument for the World Bank size rank using the US Census size-rank. The F-statistic comes from the first stage rank-rank regression of World Bank size rank on US Census size rank. In columns (7) & (8), we add ISIC 3-digit fixed effects, so that the size-rank coefficients are estimated using variation within ISIC3 over time in its country-year industry rank. The ISIC3 fixed effect models result in a drop in sample size because the ISIC3 panel structure does not exist for some of the early WBES country surveys.

TABLE 3: IMPACT OF FIRM SIZE ON TAX COMPLIANCE

	OLS		Reduced Form		IV	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A						
Outcome: Full Tax Compliance {0,1}						
Industry Size Rank (World Bank)	0.244***	0.256***			0.519***	0.401**
	(0.045)	(0.072)			(0.138)	(0.169)
Industry Size Rank (US Census)			0.175***	0.132**		
			(0.045)	(0.062)		
FE country*year*controls	✗		✗		✗	
FE country*year*ISIC2		✗		✗		✗
F-stat					35.12	38.80
R-squared	0.626	0.736	0.624	0.733	0.617	0.735
Panel B						
Outcome: % Sales Reported						
Industry Size Rank (World Bank)	0.152***	0.150***			0.221**	0.183
	(0.034)	(0.047)			(0.088)	(0.114)
Industry Size Rank (US Census)			0.075**	0.060		
			(0.035)	(0.041)		
FE country*year*controls	✗		✗		✗	
FE country*year*ISIC2		✗		✗		✗
F-stat					35.12	38.80
R-squared	0.719	0.795	0.715	0.793	0.718	0.795
Observations	3,187	2,348	3,187	2,348	3,187	2,348
Number of clusters	83	83	83	83	83	83

Source: 2001 US Census and World Bank Enterprise Surveys 2003-2007.
Robust standard errors in parentheses, clustered at the country level. *** p<0.01, ** p<0.05, * p<0.1
Table 3 shows the coefficients of regressions of industry firm-size ranking on tax compliance, discussed in section 5.1. Tax compliance is measured as the response to the survey question: "Recognizing the difficulties many enterprises face in fully complying with taxes and regulations, what percentage of sales would you estimate the typical establishment in your area of activity reports for tax purposes?". Panel A defines full tax compliance as a dummy equal to one if a firm reports reporting 100% of its sales for tax purposes. Panel B uses the percentage of sales reported for tax purpose. World Bank size rank is the ISIC 3-digit rank of number of permanent employees per firm in the WBES country-year. US Census size rank is the ISIC 3-digit rank of number of employees per establishment in 2001, where we exclude establishments with less than 5 employees. Odd number columns include interactive fixed effects between country, year and 3 dummies indicating whether the ISIC3 industry was above median in the country-year distributions of age, export-share, and foreign-industry share. Even number columns include interactive fixed effects between country, year and 2-digit ISIC codes. In this case, size-rank coefficients are estimated using variation between 3-digit ISIC codes, controlling non-parametrically for country-year-2digit ISIC effects. For a given specification, the sample size in the even columns is smaller than in the odd columns. This is because the ISIC2 interactive fixed effect specification drops all cells where there does not exist more than one ISIC3-country-year observation within a ISIC2-country-year observation. The sample drop occurs in industries that are less specialized, which are more prevalent in less developed countries. In columns (5) & (6), we instrument for the World Bank size rank using the US Census size-rank. The F-statistic comes from the first stage rank-rank regression of World Bank size rank on US Census size rank.The tax compliance question was only administered in the earlier waves of the WBES surveys which were not structured as panels. Hence, we cannot esimate panel models for the tax compliance outcomes.

TABLE 4: IMPACT OF FIRM SIZE ON INFORMAL TAX PAYMENTS

	OLS		Reduced Form		IV		Panel: OLS	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A								
Outcome: Any Informal Payment? {0,1}								
Industry Size Rank (World Bank)	0.011	-0.024			-0.048	-0.019	-0.041	-0.046
	(0.015)	(0.017)			(0.032)	(0.073)	(0.041)	(0.058)
Industry Size Rank (US Census)			-0.015	-0.008				
			(0.010)	(0.017)				
FE country*year*controls	✗		✗		✗		✗	
FE country*year*ISIC2		✗		✗		✗		✗
FE ISIC3							✗	✗
F-stat					184.39	116.58		
R-squared	0.745	0.825	0.745	0.825	0.744	0.896	0.912	0.924
Panel B								
Outcome: % Sales Paid Informally								
Industry Size Rank (World Bank)	-0.0036	-0.0072***			-0.0114**	-0.0135	-0.0198***	-0.0221***
	(0.0024)	(0.0025)			(0.0047)	(0.0089)	(0.0059)	(0.0075)
Industry Size Rank (US Census)			-0.0036**	-0.0036				
			(0.0015)	(0.0024)				
FE country*year*controls	✗		✗		✗		✗	
FE country*year*ISIC2		✗		✗		✗		✗
FE ISIC3							✗	✗
F-stat					184.39	116.58		
R-squared	0.466	0.604	0.466	0.604	0.464	0.604	0.778	0.777
Observations	9,538	7,261	9,538	7,261	9,538	7,261	4,381	3,390
Number of clusters	137	137	137	137	137	137	74	78

Source: US Census 2001 and World Bank Enterprise Surveys 2003-2015.
Robust standard errors in parentheses, clustered at the country level. *** p<0.01, ** p<0.05, * p<0.1
Table 4 shows the coefficients of regressions of industry firm-size ranking on informal tax payments, discussed in section 5.1. Informal tax payments are measured as the answer to the question: "We've heard that establishments are sometimes required to make gifts or informal payments to public officials to 'get things done' with regards to customs, taxes, licenses, regulations, services,etc. On average, what percent of total annual sales do establishments like this one pay in informal payments or gifts to public officials for tihs purpose?". Panel A defines informal tax payment as a dummy equal to one if a firm reports paying any strictly positive amount of informal payments. Panel B uses the percentage of sales paid in informal payments. World Bank size rank is the ISIC 3-digit ranking of number of permanent employees per firm in the WBES country-year. US Census size rank is the ISIC 3-digit rank of number of employees per establishment in 2001, where we exclude establishments with less than 5 employees. Odd number columns include interactive fixed effects between country, year and three dummies indicating whether the ISIC3 industry was above median in the country-year distributions of age, export-share, and foreign-industry share. Even number columns include interactive fixed effects between country, year and 2-digit ISIC codes. For a given specification, the sample size in the even columns is smaller than in the odd columns. This is because the ISIC2 interactive fixed effect specification drops all cells where there does not exist more than one ISIC3-country-year observation within a ISIC2-country-year observation. The sample drop occurs in industries that are less specialized, which are more prevalent in less developed countries. In columns (5) & (6), we instrument for the World Bank size rank using the US Census size-rank. The F-statistic comes from the first stage rank-rank regression of World Bank size rank on US Census size rank. In columns (7) & (8), we add ISIC 3-digit fixed effects, so that the size-rank coefficients are estimated using variation within ISIC3 over time in its country-year industry rank. The ISIC3 fixed effect models result in a drop in sample size because some countries do not have repeated surveys.

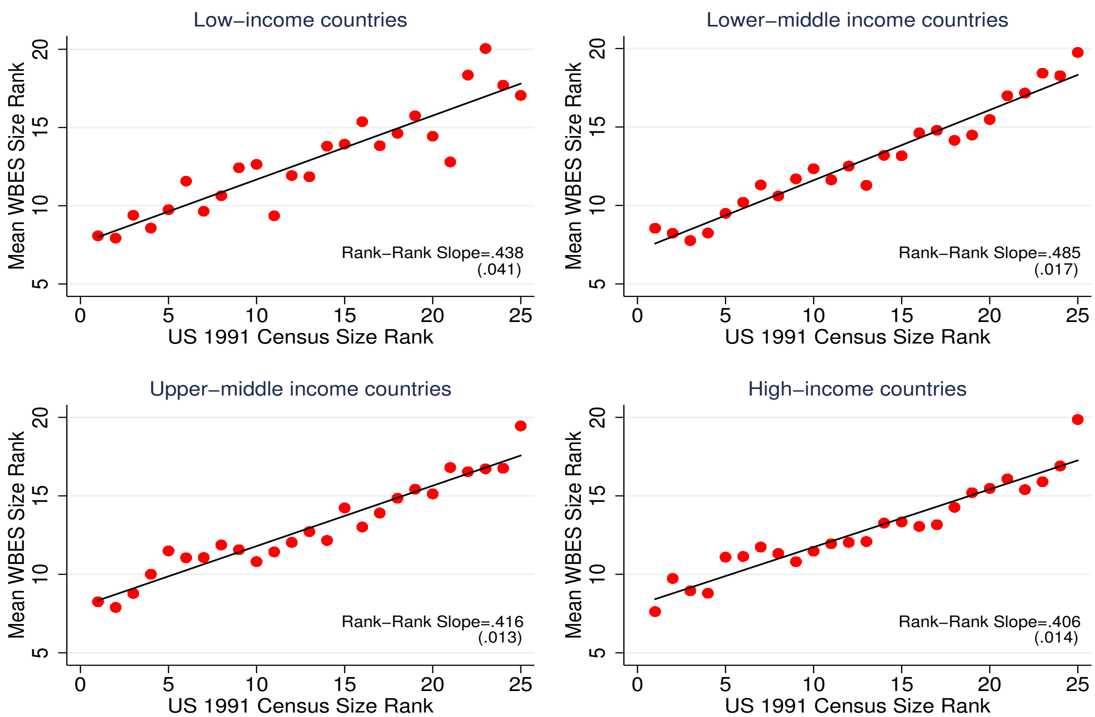
TABLE 5: IMPACT OF FIRM SIZE ON FORMAL AND TOTAL EFFECTIVE TAX RATES

	OLS		Reduced Form		IV	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A						
Outcome: Formal Effective Tax Rate						
= Tax Rate * Share of Sales Reported						
Industry Size Rank (World Bank)	0.065***	0.069***			0.112***	0.100**
	(0.015)	(0.019)			(0.035)	(0.045)
Industry Size Rank (US Census)			0.038**	0.034**		
			(0.015)	(0.017)		
FE country*year*controls	✗		✗		✗	
FE country*year*ISIC2		✗		✗		✗
F-Stat					33.99	38.73
R-squared	0.719	0.795	0.715	0.793	0.718	0.795
Observations	2,920	2,149	2,920	2,149	2,920	2,149
Number of clusters	74	74	74	74	74	74
Panel B						
Outcome: Total Effective Tax Rate						
= Sum of Formal and Informal						
Industry Size Rank (World Bank)	0.064***	0.069***			0.033	0.024
	(0.023)	(0.022)			(0.059)	(0.075)
Industry Size Rank (US Census)			0.011	0.008		
			(0.021)	(0.026)		
FE country*year*controls	✗		✗		✗	
FE country*year*ISIC2		✗		✗		✗
F-stat					33.38	38.73
R-squared	0.753	0.811	0.750	0.809	0.752	0.810
Observations	2,731	1,970	2,731	1,970	2,731	1,970
Number of clusters	74	74	74	74	74	74

Source: 2001 US Census and World Bank Enterprise Surveys 2003-2007 and tax rates collected by the authors based on the KPMG's corporate and indirect tax rates tables.
Robust standard errors in parentheses, clustered at the country level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$
Table 5 shows the coefficients of regressions of industry firm-size ranking on the formal effective tax rate (Panel A) and the total effective tax rate (Panel B), discussed in Section 5.3. In Panel A, the effective formal tax rate is calculated as the product of the (firm-specific) share of sales reported for tax purpose (outcome in Table 3) and the (country-yea specific) sum of the indirect and corporate tax rates. In Panel B, the total effective tax rate is calculated, at the firm level, as the sum of formal tax payments and informal tax payments (outcome in Table 4) as a share of a firms' sales.World Bank size rank is the ISIC 3-digit rank of number of permanent employees per firm in the WBES country-year. US Census size rank is the ISIC 3-digit rank of number of employees per establishment in 2001, where we exclude establishments with less than 5 employees. Odd number columns include interactive fixed effects between country, year and 3 dummies indicating whether the ISIC3 industry was above median in the country-year distributions of age, export-share, and foreign-industry share. Even number columns include interactive fixed effects between country, year and 2-digit ISIC codes. In this case, size-rank coefficients are estimated using variation between 3-digit ISIC codes, controlling non-parametrically for country-year-2digit ISIC effects. For a given specification, the sample size in the even columns is smaller than in the odd columns. This is because the ISIC2 interactive fixed effect specification drops all cells where there does not exist more than one ISIC3-country-year observation within a ISIC2-country-year observation. The sample drop occurs in industries that are less specialized, which are more prevalent in less developed countries. In columns (5) & (6), we instrument for the World Bank size rank using the US Census size-rank. The F-statistic comes from the first stage rank-rank regression of World Bank size rank on US Census size rank. The question on formal tax payments was only administered in the earlier waves of the WBES surveys which were not structured as panels. Hence, we can not estimate panel models for the formal and total effective tax rate outcomes.

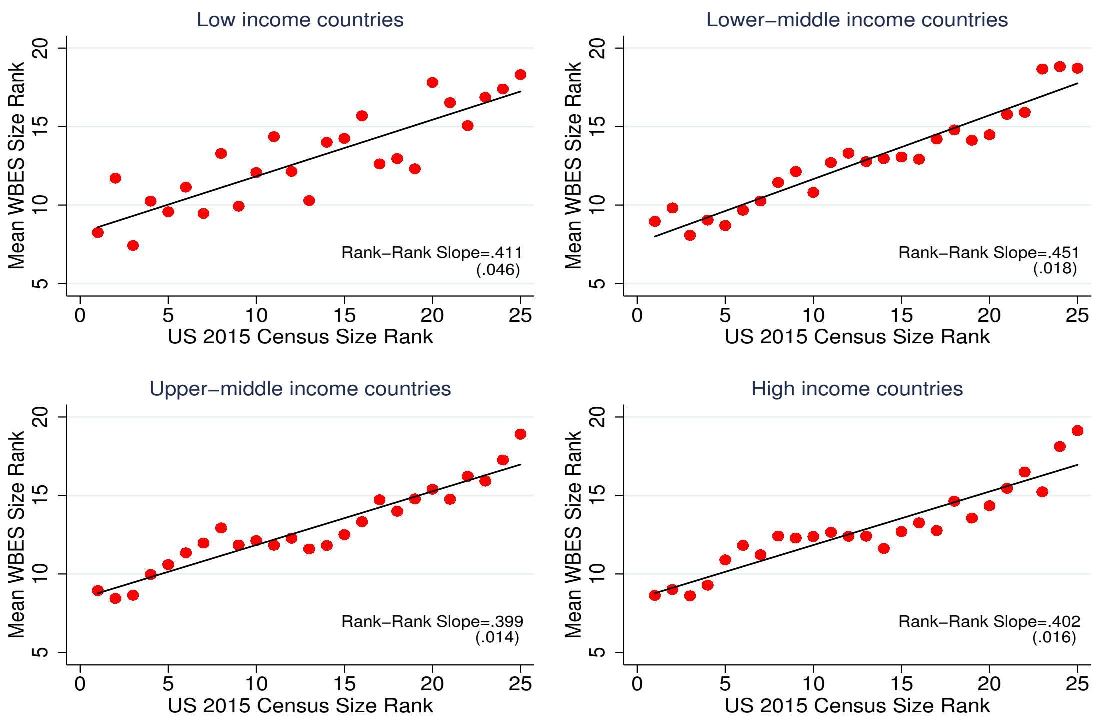
Appendix Figures & Tables (Not for Publication)

FIGURE A1: INDUSTRY SIZE IN US CENSUS 91 ON INDUSTRY SIZE IN WBES BY INCOME



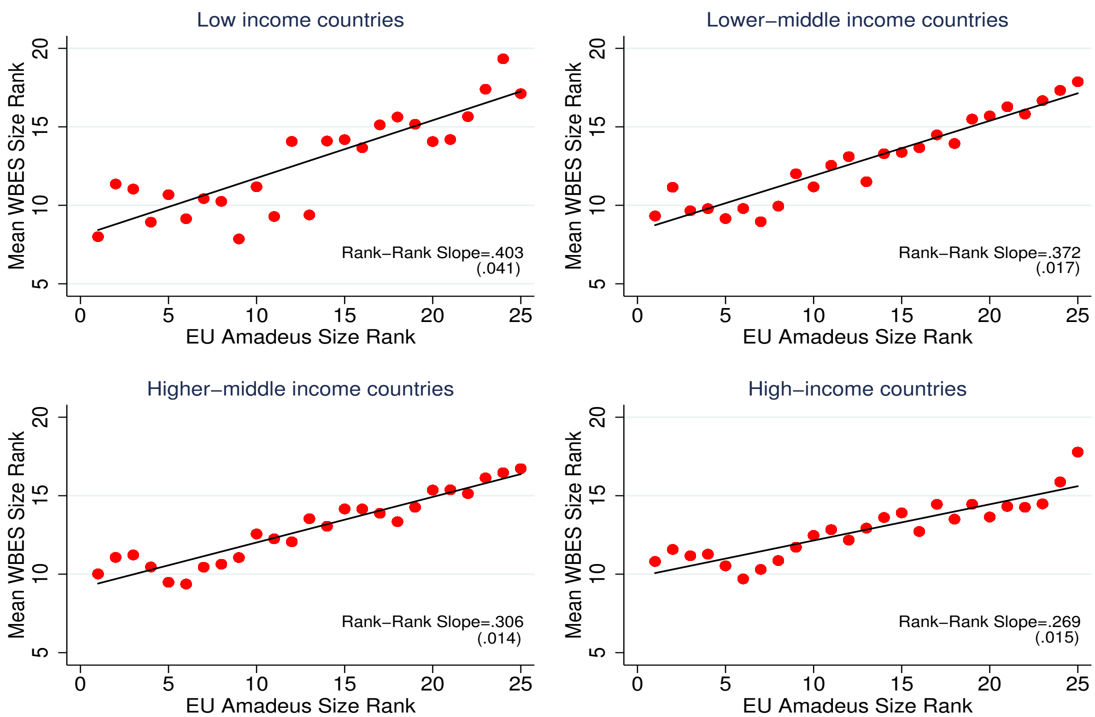
Source: 1991 US Census and World Bank Enterprise Surveys 2003-2015.
Figure A1 presents non-parametric binned scatterplot of the relationship between WBES countries' and US' 1991 percentile firm size ranks in the four per capita income groups based on the World Bank classification. The USD cut-off values used for the classification are: 1100 for low income countries; between 1100 and 3900 for lower-middle income countries; between 3900 and 12730 for upper-middle income countries; above 12730 for high income countries. This figure is based on the full WBES sample and the US Census 1991. WBES size rank is the ISIC3 rank of number of permanent employees per firm in the WBES country-year. US Census size rank is the ISIC3 rank of number of employees per establishment in 199, where we exclude establishments with less than 5 employees. Within each per capita income group, the WBES and US Census ranks are then grouped into 25 equal sized bins. For each per capita income group, this graph plots the mean WBES size rank percentile within each 25-quantile US size rank percentile, together with a best-fit line. The rank-rank slope coefficient and standard error are estimated in each per capita income group using the underlying ISIC3-year-country data.

FIGURE A2: INDUSTRY SIZE IN US CENSUS 15 ON INDUSTRY SIZE IN WBES BY INCOME



Source: 2015 US Census and World Bank Enterprise Surveys 2003-2015.
Figure A2 presents non-parametric binned scatterplot of the relationship between WBES countries' and US 2015 percentile firm size ranks in the four per capita income groups based on the World Bank classification. The USD cut-off values used for the classification are: 1100 for low income countries; between 1100 and 3900 for lower-middle income countries; between 3900 and 12730 for upper-middle income countries; above 12730 for high income countries. This figure is based on the full WBES sample and the US Census 2015. WBES size rank is the ISIC3 rank of number of permanent employees per firm in the WBES country-year. US Census size rank is the ISIC3 rank of number of employees per establishment in 2015, where we exclude establishments with less than 5 employees. Within each per capita income group, the WBES and US Census ranks are then grouped into 25 equal sized bins. For each per capita income group, this graph plots the mean WBES size rank percentile within each 25-quantile US size rank percentile, together with a best-fit line. The rank-rank slope coefficient and standard error are estimated in each per capita income group using the underlying ISIC3-year-country data

FIGURE A3: INDUSTRY SIZE IN EU AMADEUS 07 ON INDUSTRY SIZE IN WBES BY INCOME



Source: 2007 EU Amadeus and World Bank Enterprise Surveys 2003-2015.
Figure A3 presents non-parametric binned scatterplot of the relationship between WBES countries' and EU' percentile firm size ranks in the four per capita income groups based on the World Bank classification. The USD cut-off values used for the classification are: 1100 for low income countries; between 1100 and 3900 for lower-middle income countries; between 3900 and 12730 for upper-middle income countries; above 12730 for high income countries. This figure is based on the full WBES sample and the EU Amadeus sample of UK and German firms in 2007. WBES size rank is the ISIC3 rank of number of permanent employees per firm in the WBES country-year. EU Amadeus Census size rank is the ISIC3 rank of number of employees per establishment in 2007 in Germany and UK. Within each per capita income group, the WBES and EU Amadeus ranks are then grouped into 25 equal sized bins. For each per capita income group, this graph plots the mean WBES size rank percentile within each 25-quantile EU Amadeus size rank percentile, together with a best-fit line. The rank-rank slope coefficient and standard error are estimated in each per capita income group using the underlying ISIC3-year-country data.

TABLE A1: FIRST-STAGE AND IV FOR DIFFERENT INSTRUMENTAL VARIABLES

Panel A: First Stage	Industry Size Rank (WBES)					
	(1)	(2)	(3)	(4)	(5)	(6)
Industry Size Rank (US Census 01)	0.308***					
	(0.019)					
Industry Size Rank (US Census 91)		0.337***				
		(0.022)				
Industry Size Rank (US Census 15)			0.253***			
			(0.020)			
Industry Size Rank (EU Amadeus 07)				0.101***		
				(0.017)		
Industry Size Rank (KRZ weighting)					0.082***	
					(0.031)	
Industry Size Mean (US Census 01)						0.342***
						(0.048)
R-squared	0.552	0.562	0.524	0.501	0.624	0.140
Panel B: IV Results	Tax Inspection {0,1}					
	(1)	(2)	(3)	(4)	(5)	(6)
Industry Size Rank (WBES)	0.242***					
	(0.053)					
Industry Size Rank (WBES)		0.188***				
		(0.048)				
Industry Size Rank (WBES)			0.167***			
			(0.050)			
Industry Size Rank (WBES)				0.765***		
				(0.217)		
Industry Size Rank (WBES)					0.492***	
					(0.152)	
Industry Size Mean (WBES)						0.091***
						(0.020)
1st stage rank-rank based on sample	USCen01	USCen91	USCen15	EUAma07	USCen01RKZ	USCen01
1st stage F-statistic	276.83	227.40	159.62	35.52	6.94	50.87
R-squared	0.662	0.662	0.659	0.584	0.628	0.554
FE country*year*controls	✗	✗	✗	✗	✗	✗
Observations	11,675	11,671	11,031	11,743	11,710	11,675
Number of clusters	140	140	140	140	140	140

Sources: World Bank Enterprise Surveys 2003-2015 and US firm census 2001.
Robust standard errors in parentheses, clustered at the country level. *** p<0.01, ** p<0.05, * p<0.1
Table A1 shows first stage regressions and IV regressions on tax inspections across different samples of first stage industry firm-size. All regressions include interactive fixed effects between country, year and 3 dummies indicating whether the ISIC3 industry was above median in the country-year distributions of age, export-share, and foreign-industry share. All regressions Industry size rank is constructed similarly to the main tables: please see Table 1 for more detail. In Panel A row 1-3, US industry size is constructed using the 2001, 1991, and 2015 US Census of firm employment. In Panel A row 4, the industry size is constructed using the Amadeus sample in the UK and Germany between 1998 and 2007. In Panel A row 5, the industry average firm-size is constructed using the Kumar et al. [1999] 'employee-weighted' methodology (KRZ). This methodology effectively places more weight on firms that concentrate large shares of total employment when calculating the industry-wide average firm size. In Panel A row 6, the industry size mean (as opposed to rank) in the US Census 2001 is used as the regressor. In Panel A Cols.1-4, the outcome variable is the WBES size rank as defined in the main tables. In Panel A Column 5, the outcome variable is the average firm size in all WBES countries calculated using the employee-weighted KRZ methodology. In Panel A Column 6, the outcome variable is the industry mean firm size (as opposed to rank). In Panel B, the outcome variable is a dummy taking a value of one if the firm has had any tax inspections over the past year: please see Table 1 for more details. In Panel B, Columns 1-5, the WBES industry size rank is instrumented for using a size rank constructed from different samples, respectively: US Census 2001, US Census 1991, US Census 2015, EU Amadeus, US Census 2001 RKZ reweigthed. In Panel B Column 6, the WBES industry size mean is instrumented for using the US Census 2001 industry size mean.

TABLE A2: IMPACT OF FIRM SIZE ON TAX INSPECTION, RESTRICTED TO MANUFACTURING

Outcome:	Industry Size Rank (WBES)	Tax Inspection {0,1}		# Inspections	
	(1)	(2)	(3)	(4)	(5)
	1st Stage	Red. Form	IV	Red. Form	IV
Industry Size Rank (US Census)	0.310***	0.135***		0.009***	
	(0.029)	(0.034)		(0.002)	
Industry Size Rank (WBES)			0.434***		0.0289***
			(0.089)		(0.005)
1st stage F-statistic			114.36		114.36
FE country*year*controls	✗	✗	✗	✗	✗
R-squared	0.667	0.685	0.688	0.614	0.617
Observations	6,726	6,726	6,726	6,726	6,726
Number of clusters	138	138	138	138	138

Sources: World Bank Enterprise Surveys 2003-2015 and US firm Census 2001.
Robust standard errors in parentheses, clustered at the country level. *** p<0.01, ** p<0.05, * p<0.1
Table A2 shows the coefficients of regressions of industry firm-size ranking on various outcomes, limiting the sample to manufacturing industries. All regressions include interactive fixed effects between country, year and 3 dummies indicating whether the ISIC3 industry was above median in the country-year distributions of age, export-share, and foreign-industry share. Manufacturing is defined as the set of ISIC2 industries which lie between code-number 15 and 37, using the ISIC Classification version 3.1. Size rank is constructed similarly to the main tables: please see Table 1 for further details. In Column 1, the outcome is WBES industry size rank. In Columns 2-3, the outcome variable is constructed similarly to Table 1. In Columns 4-5, the outcome variable is constructed similarly to Table 2. In Columns 3 and 5, the WBES industry size rank is instrumented for using the US industry size rank.

TABLE A3: TAX INSPECTION ON FIRM SIZE AND CAPITAL NEEDS

Outcome:	Tax Inspection {0,1}		# Inspections	
	(1)	(2)	(3)	(4)
Industry Size Rank (US Census)	0.063***	0.090***	0.003***	0.006***
	(0.020)	(0.032)	(0.001)	(0.002)
Industry External Reliance Rank (US Compustat)	-0.001		-0.001	
	(0.023)		(0.001)	
Industry Capital-Labor Ratio Rank (US Compustat)		-0.001		0.002
		(0.027)		(0.002)
FE country*year*controls	✗	✗	✗	✗
R-squared	0.580	0.624	0.534	0.562
Observations	9,715	5,320	9,715	5,320
Number of clusters	131	106	131	106

Sources: World Bank Enterprise Surveys 2003-2015 and US firm census 2001.
Robust standard errors in parentheses, clustered at the country level. *** p<0.01, ** p<0.05, * p<0.1
Table A3 shows the coefficients of regressions of industry size ranking and industry capital ranking on various outcomes. All regressions include interactive fixed effects between country, year and three dummies indicating whether the ISIC3 industry was above median in the country-year distributions of age, export-share, and foreign-industry share. Columns 1 and 3 include the industry external reliance rank as an additional explanatory variable. External reliance is defined as the share of capital expenditure not financed by cash flow of operations. This measure can directly be constructed for a subsample of firms in WBES. In the US, we rely on Compustat firm-level data between 1997 and 2007. We follow the methodology of Rajan and Zingales [1998] to construct the ISIC3 measure of external reliance. Having constructed the external reliance at the firm level in both the US and WBES, we construct ISIC3-country-year rankings similar to the industry employee-size rankings. Columns 2 and 4 include the industry capital-labor ratio rank as an additional explanatory variable. The capital-labor ratio is defined as the ratio of assets to number of permanent employees. This measure can be constructed for a subsample of firms in WBES. In the US, we rely on Compustat firm-level data between 1997 and 2007. The drop in sample size in columns 2 and 4 is due to the missing observations on assets for firms in the WBES surveys.

TABLE A4: SIZE-GRADIENTS INTERACTION WITH STATUTORY TAX RATES

Outcome:	(1) Tax Inspection {0,1}	(2) # Inspections	(3) Tax Compliance	(4) Informal Tax
Industry Size Rank (US Census)	0.041*	0.003*	-0.003	-0.003
	(0.024)	(0.002)	(0.053)	(0.002)
Industry Size Rank*1(High Tax Rate)	0.061**	0.003	0.129**	-0.002
	(0.030)	(0.002)	(0.062)	(0.003)
FE country*year*controls	✗	✗	✗	✗
F-test joint significance [Rank + Rank*1(High Rate)]	32.66	57.08	4.43	14.93
P-value	0.0000	0.0000	0.0373	0.0002
R-squared	0.649	0.579	0.712	0.468
Observations	11,103	11,103	2,920	9,033
Number of clusters	127	127	125	74

Sources: World Bank Enterprise Surveys 2003-2015 and US firm Census 2001.
Table A4 shows the coefficients of regressions of industry firm-size ranking, and its interaction with statutory tax rate, on tax inspection, formal compliance, and informal tax payments. US Census size rank is the ISIC 3-digit rank of number of employees per establishment in 2001, where we exclude establishments with less than 5 employees. All regressions include interactive fixed effects between country, year and 3 dummies indicating whether the ISIC3 industry was above median in the country-year distributions of age, export-share, and foreign-industry share. In addition, all regressions include an interaction between the Census size rank and a country-year varying dummy which takes a value of 1 if the WBES country-year statutory rate was above the median in the WBES sample distribution. The statutory rate is constructed as the sum of the corporate tax rate and the indirect tax rate. The F-test value and p-value report the results from an F-test on the joint significance of the industry size rank coefficient and the interaction coefficient between size rank and the high tax rate dummy.

TABLE A5: CORRELATION OF SIZE WITH PERFORMANCE

Outcome:	(1) Log Sales	(2) Cost per Employee	(3) Labor regulations	(4) Intl. Certificate	(5) Part of Larger Firm	(6) Formal Training	(7) Finance Constraints
Industry Size Rank (US Census)	0.011***	0.011***	0.002***	0.003***	0.001**	0.001***	-0.001***
	(0.001)	(0.002)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
FE country*year*controls	✗	✗	✗	✗	✗	✗	✗
R-squared	0.892	0.896	0.516	0.456	0.580	0.624	0.213
Observations	9,415	9,175	8,160	9,544	9,553	9,691	9,705
Number of clusters	131	130	129	131	131	130	131

Sources: World Bank Enterprise Surveys 2003-2015 and US firm census 2001.
Robust standard errors in parentheses, clustered at the country-year level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$
Table A5 shows the coefficients of regressions of industry size ranking on various outcomes. All regressions include interactive fixed effects between country, year and 3 dummies indicating whether the ISIC3 industry was above median in the country-year distributions of age, export-share, and foreign-industry share. Industry size rank is constructed similarly to the main tables: please see Table 1 for further details. In Column 1, the outcome is the natural logarithm of annual firm sales. In Column 2, the outcome is the ratio of annual full-time permanent employee labor costs to the total number of full-time permanent employees. In Column 3, the outcome variable is the response to the answer "are labor regulations no/minor/major/very severe obstacle to the current operations of this establishment?" (higher value meaning more of an obstacle). In Column 4, the outcome variable is a dummy taking value one if the establishment reports having an internationally-recognized quality certification. In Column 5, the outcome variable is a dummy taking value one if the establishment is part of a larger firm. In Column 6, the outcome variable is a dummy taking value one if the establishment reports having formal training courses for its permanent employees. In Column 7, the outcome variable is the response to the answer "is access to finance (availability and cost) no/minor/major/very severe obstacle to the current operations of this establishment?" (higher value meaning more of an obstacle).